Gerhard Krauss

Biochemistry of Signal Transduction and Regulation

Third, Completely Revised Edition

WILEY-VCH Verlag GmbH & Co. KGaA

VISIT…

Prof. Dr. Gerhard Krauss
Laboratorium für Biochemie
Universität Bayreuth
95440 Bayreuth
Germany
Gerhard.Krauss@uni-bayreuth.de

This book was carefully produced. Nevertheless, author and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

1st edition 1999
2nd edition 2001
3rd edition 2003

Cover illustration by Hanno Krauss, Bayreuth

Library of Congress Card No.:
Applied for.

British Library Cataloguing-in-Publication Data:
A catalogue record for this book is available from the British Library.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>

© 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Printed in the Federal Republic of Germany

Printed on acid-free paper

Composition Mitterweger & Partner, Kommunikationsgesellschaft mbH, Plankstadt
Printing Druckhaus Darmstadt GmbH, Darmstadt
Bookbinding Litges & Dopf Buchbinderei GmbH, Heppenheim
ISBN 3-527-30591-2

Preface

This book has originated from lectures on regulation and signal transduction that are offered to students of biochemistry, biology and chemistry at the University of Bayreuth. The idea to write a book on signal transduction was born during the preparations of these lectures where I realized that it is extremely difficult to achieve an overview of the area of signal transduction and regulation and to follow the progress of this field. The first book appeared in 1997 and was written in German. It was soon substituted by two successive English editions that are now followed by the 3rd edition which includes data and references up to 2002.

Cellular signaling in higher organisms is a major topic in modern medical and pharmacological research and is of central importance in the biomolecular sciences. Accordingly, the book concentrates on signaling and regulation in animal systems and in man. Plant systems could not be considered, and results from lower eukaryotes and prokaryotes are only cited if they are of exemplary character. The enormous increase in data on signal transduction has led me to leave out the chapter on ion channels and nerve signaling found in the former editions. This topic has since evolved into a huge research area of its own that could not be considered adequately within this book.

Our knowledge of signal transduction processes has exploded in the past 10 to 15 years, and the basic principles of intra- and intercellular signaling are now quite well established. Signaling processes can be described nowadays more and more on a molecular level and structure-function relationships of many central signaling proteins have been worked out. Research on signal transduction is presently focused on the characterization of the distinct cellular functions of the huge number of different signaling proteins and their subspecies, on the supramolecular organization of signaling proteins and on the interplay between different signaling pathways. The enormous complexity of signaling systems revealed by these studies makes it increasingly difficult to write a book that provides a truly comprehensive overview on signal transduction and considers all of the major new achievements. In consequence, not all branches and fields of signal transduction could be treated here with the same thoroughness.

It is the aim of the present book to describe the structural and biochemical properties of signaling molecules and their regulation, the interaction of signaling proteins at

the various levels of signal transduction and to work out the basic principles of cellular communication. Numerous studies in very diverse systems have revealed that the basic principles of signaling and regulation are similar in all higher organisms. Therefore, the book concentrates on the best studied reactions and components of selected signaling pathways and does not attempt to describe distinct signaling pathways (e.g. the vision process) in its entirety. Furthermore, results from very different eucaryotic organisms and tissues have been included. Due to the huge number of publications on the topic, mostly review articles are cited. Only a few original articles have been selected on a more or less subjective basis.

I am grateful to all people who have encouraged me to continue with the book and who have supported me with many helpful comments and corrections. In first place I want to thank my colleague Mathias Sprinzl and my former coworkers Thomas Hey, Carl Christian Gallert and Oliver Hobert. I am also grateful to Hannes Krauss and Yiwei Huang for the figures and structure representations.

Bayreuth, June 2003 Gerhard Krauss

Contents

Preface *VI*

1 The Regulation of Gene Expression *1*
1.1 Regulation of Gene Expression: How and Where? A Schematic Overview *1*
1.2 Protein-Nucleic Acid Interactions as a Basis for Specific Gene Regulation *3*
1.2.1 Structural Motifs of DNA-binding Proteins *3*
1.2.2 The Nature of the Specific Interactions in Protein-Nucleic Acid Complexes *9*
1.2.3 The Role of the DNA Conformation in Protein-DNA Interactions *11*
1.2.4 Structure of the Recognition Sequence and Quaternary Structure of DNA-binding Proteins *13*
1.3 The Principles of Transcription Regulation *17*
1.3.1 Elements of Transcription Regulation *17*
1.3.2 Functional Requirements for Repressors and Transcriptional Activators *19*
1.3.3 Mechanisms for the Control of the Activity of DNA-binding Proteins *20*
1.3.3.1 Binding of Effector Molecules *21*
1.3.3.2 Binding of Inhibitory Proteins *23*
1.3.3.3 Modification of Regulatory Proteins *23*
1.3.3.4 Changes in the Concentration of Regulatory DNA-binding Proteins *24*
1.4 Regulation of Transcription in Eucaryotes *25*
1.4.1 Overview of Transcription Initiation in Procaryotes *26*
1.4.2 The Basic Features of Eukaryotic Transcription *28*
1.4.3 The Eucaryotic Transcription Apparatus *30*
1.4.3.1 Structure of the Transcription Start Site and Regulatory Sequences *30*
1.4.3.2 Elementary Steps of Eucaryotic Transcription *32*
1.4.3.3 Formation of a Basal Transcription Apparatus from General Transcription Factors and RNA Polymerase *33*
1.4.3.4 Phosphorylation of RNA Polymerase II and the Onset of Transcription *36*
1.4.3.5 TFIIH – a Pivotal Regulatory Protein Complex *38*

1.4.4 Regulation of Eucaryotic Transcription by DNA-binding Proteins *39*
1.4.4.1 The Structure of Eucaryotic Transcriptional Activators *39*
1.4.4.2 Concerted Action of Transcriptional Activators and Coactivators in the Regulation of Transcription *41*
1.4.4.3 Interactions with the Transcription Apparatus *45*
1.4.5 Regulation of the Activity of Transcriptional Activators *45*
1.4.5.1 The Principal Pathways for the Regulation of Transcriptional Activators *46*
1.4.5.2 Phosphorylation of Transcriptional Activators *46*
1.4.5.3 Heterotypic Dimerization *50*
1.4.5.4 Regulation by Binding of Effector Molecules *52*
1.4.6 Specific Repression of Transcription *52*
1.4.7 Chromatin Structure and Transcription Activation *55*
14.7.1 Transcriptional Activity and Histone Acetylation *58*
1.4.7.2 Transcriptional Activity and Histone Methylation *62*
1.4.7.3 Enhanceosomes *63*
1.4.8 Methylation of DNA *65*
1.5 Post-transcriptional Regulation of Gene Expression *68*
1.5.1 Modifications at the 5' and 3' Ends of the Pre-mRNA *69*
1.5.2 Formation of Alternative mRNA by Alternative Polyadenylation and by Alternative Splicing *70*
1.5.3 Regulation via Transport and Splicing of Pre-mRNA *73*
1.5.4 Stability of the mRNA *75*
1.5.5 Regulation at the Level of Translation *78*
1.5.5.1 Regulation by binding of protein to the 5' end of the mRNA *79*
1.5.5.2 Regulation by Modification of Initiation Factors *80*

2 The Regulation of Enzyme Activity *89*
2.1 Enzymes as Catalysts *90*
2.2 Regulation of Enzymes by Effector Molecules *91*
2.3 Principal Features of Allosteric Regulation *93*
2.4 Regulation of Enzyme Activity by Binding of Inhibitor and Activator Proteins *94*
2.5 Regulation of Enzyme Activity by Phosphorylation *95*
2.5.1 Regulation of Glycogen Phosphorylase by Phosphorylation *97*
2.5.2 Regulation of Isocitrate Dehydrogenase (E. coli) by Phosphorylation *100*
2.6 Regulation via the Ubiquitin-Proteasome Pathway *101*
2.6.1 Components of the Ubiquitin System *102*
2.6.2 Degradation in the Proteasome *107*
2.6.3 Recognition of the Substrate in the Ubiquitin-Proteasome Degradation Pathway *108*
2.6.4 Regulatory Function of Ubiquitin Conjugation and the Targeted Degradation of Proteins *110*
2.7 Regulation of Proteins by Sumoylation *113*

3 Structure and Function of Signal Pathways *115*
3.1 General Function of Signal Pathways *115*
3.2 Structure of Signaling Pathways *117*
3.2.1 The Mechanisms of Intercellular Communication *117*
3.2.2 Principles of Intracellular Signal Transduction *119*
3.2.3 Components of Intracellular Signal Transduction *120*
3.2.4 Coupling of Proteins in Signaling Chains *122*
3.2.4.1 Coupling by Specific Protein–Protein Interactions *122*
3.2.4.2 Coupling by Protein Modules *122*
3.2.4.3 Coupling by Reversible Docking Sites *123*
3.2.4.4 Coupling by Colocalization *123*
3.2.4.5 Linearity, Branching and Crosstalk *124*
3.2.4.6 Variability and Specificity of Receptors and Signal Responses *126*
3.3 Extracellular Signaling Molecules *128*
3.3.1 The Chemical Nature of Hormones *128*
3.3.2 Hormone Analogs: Agonists and Antagonists *131*
3.3.3 Endocrine, Paracrine and Autocrine Signaling *133*
3.3.4 Direct Modification of Protein by Signaling Molecules *133*
3.4 Hormone Receptors *135*
3.4.1 Recognition of Hormones by Receptors *135*
3.4.2 The Interaction between Hormone and Receptor *135*
3.5 Signal Amplification *139*
3.6 Regulation of Inter- and Intracellular Signaling *141*
3.7 Membrane Anchoring and Signal Transduction *142*
3.7.1 Myristoylation *144*
3.7.2 Palmitoylation *145*
3.7.3 Farnesylation and Geranylation *146*
3.7.4 The Glycosyl-Phosphatidyl-Inositol Anchor (GPI Anchor) *147*
3.7.5 The Switch Function of Lipid Anchors *148*

4 Signaling by Nuclear Receptors *151*
4.1 Ligands of Nuclear Receptors *151*
4.2 Principles of Signaling by Nuclear Receptors *153*
4.3 Classification and Structure of Nuclear Receptors *156*
4.3.1 DNA-Binding Elements of Nuclear Receptors, HREs *156*
4.3.2 The DNA-Binding Domain of Nuclear Receptors *159*
4.3.3 HRE Recognition and Structure of the HRE-Receptor Complex *161*
4.3.4 Ligand-binding Domains *162*
4.3.5 Transactivating Elements of the Nuclear Receptors *164*
4.4 Mechanisms of Transcriptional Regulation by Nuclear Receptors *165*
4.5 Regulation and Variability of Signaling by Nuclear Receptors *169*
4.6 The Signaling Pathway of the Steroid Hormone Receptors *171*
4.7 Signaling by Retinoids, Vitamin D3, and the T3-Hormone *173*
4.7.1 Structure of the HREs of RXR Heterodimers *175*
4.7.2 Complexity of the Interaction between HRE, Receptor and Hormone *175*

5 G Protein-Coupled Signal Transmission Pathways *179*
5.1 Transmembrane Receptors: General Structure and Classification *179*
5.2 Structural Principles of Transmembrane Receptors *181*
5.2.1 The Extracellular Domain of Transmembrane Receptors *181*
5.2.2 The Transmembrane Domain *183*
5.2.3 The Intracellular Domain of Membrane Receptors *185*
5.2.4 Regulation of Receptor Activity *186*
5.3 G Protein-Coupled Receptors *187*
5.3.1 Structure of G Protein-Coupled Receptors *188*
5.3.2 Ligand Binding *191*
5.3.3 Mechanism of Signal Transmission *192*
5.3.4 Switching Off and Desensitization of 7-Helix Transmembrane Receptors *192*
5.3.5 Dimerization of GPCRs *196*
5.4 Regulatory GTPases *197*
5.4.1 The GTPase Superfamily: General Functions and Mechanism *197*
5.4.2 Inhibition of GTPases by GTP Analogs *200*
5.4.3 The G-domain as Common Structural Element of the GTPases *200*
5.4.4 The Different GTPase Families *201*
5.5 The Heterotrimeric G Proteins *202*
5.5.1 Classification of the Heterotrimeric G Proteins *203*
5.5.2 Toxins as Tools in the Characterization of Heterotrimeric G Proteins *205*
5.5.3 The Functional Cycle of Heterotrimeric G Proteins *206*
5.5.4 Structural and Mechanistic Aspects of the Switch Function of G Proteins *208*
5.5.5 Structure and Function of the βγ-Complex *215*
5.5.6 Membrane Association of the G Proteins *217*
5.5.7 Regulators of G Proteins: Phosducin and RGS Proteins *218*
5.6 Effector Molecules of G Proteins *220*
5.6.1 Adenylyl Cyclase and cAMP as Second Messenger *220*
5.6.2 Phospholipase C *225*

6 Intracellular Messenger Substances: Second Messengers *231*
6.1 General Functions of Intracellular Messenger Substances *231*
6.2 cAMP *233*
6.3 cGMP *235*
6.4 Metabolism of Inositol Phospholipids and Inositol Phosphates *237*
6.5 Inositol 1,4,5-Triphosphate and Release of Ca^{2+} *240*
6.5.1 Release of Ca^{2+} from Ca^{2+} Storage *241*
6.5.2 Influx of Ca^{2+} from the Extracellular Region *245*
6.5.3 Removal and Storage of Ca^{2+} *246*
6.5.4 Temporal and Spatial Changes in Ca^{2+} Concentration *246*
6.6 Phosphatidyl Inositol Phosphates and PI3-Kinase *248*
6.6.1 PI3-Kinases *249*
6.6.2 The Messenger Substance PtdIns(3,4,5)P_3 *250*

6.6.3 Akt Kinase and PtdIns(3,4,5)P_3 Signaling *252*
6.6.4 Functions of PtIns(4,5)P_2 *253*
6.7 Ca^{2+} as a Signal Molecule *253*
6.7.1 Calmodulin as a Ca^{2+} Receptor *256*
6.7.2 Target Proteins of Ca^{2+}/Calmodulin *257*
6.7.3 Other Ca^{2+} Receptors *258*
6.8 Diacylglycerol as a Signal Molecule *259*
6.9 Other Lipid Messengers *260*
6.10 The NO Signaling Molecule *261*
6.10.1 Reactivity and Stability of NO *262*
6.10.2 Synthesis of NO *263*
6.10.3 Physiological Functions and Attack Points of NO *264*

7 Ser/Thr-specific Protein Kinases and Protein Phosphatases *269*
7.1 Classification, Structure and Characteristics of Protein Kinases *269*
7.1.1 General Classification and Function of Protein Kinases *269*
7.1.2 Classification of Ser/Thr-specific Protein Kinases *272*
7.2 Structure and Regulation of Protein Kinases *273*
7.2.1 Main Structural Elements of Protein Kinases *274*
7.2.2 Substrate Binding and Recognition *276*
7.2.3 Control of Protein Kinase Activity *277*
7.3 Protein Kinase A *280*
7.3.1 Structure and Substrate Specificity of Protein Kinase A *280*
7.3.2 Regulation of Protein Kinase A *281*
7.4 Protein Kinase C *283*
7.4.1 Characterization and Classification *283*
7.4.2 Structure and Activation of Protein Kinase C *286*
7.4.3 Regulation of Protein Kinase C *288*
7.4.4 Functions and Substrates of Protein Kinase C *290*
7.5 Ca^{2+}/Calmodulin-dependent Protein Kinases *292*
7.5.1 Importance and General Function *292*
7.5.2 Structure and Autoregulation of CaM Kinase II *293*
7.6 Ser/Thr-specific Protein Phosphatases *296*
7.6.1 Structure and Classification of Ser/Thr Protein Phosphatases *296*
7.6.2 Regulation of Ser/Thr Protein Phosphatases *297*
7.6.3 Protein Phosphatase I, PPI *299*
7.6.4 Protein Phosphatase 2A, PP2A *301*
7.6.5 Protein Phosphatase 2B, Calcineurin *302*
7.7 Regulation of Protein Phosphorylation by Subcellular Localization *305*

8 Signal Transmission via Transmembrane Receptors with Tyrosine-Specific Protein Kinase Activity *311*
8.1 Structure and Function of Receptor Tyrosine Kinases *311*
8.1.1 General Structure and Classification *313*
8.1.2 Ligand Binding and Activation *314*

8.1.3 Structure and Activation of the Tyrosine Kinase Domain *319*
8.1.4 Effector Proteins of the Receptor Tyrosine Kinases *323*
8.1.5 Attenuation and Termination of RTK Signaling *326*
8.2 Protein Modules as Coupling Elements of Signal Proteins *328*
8.2.1 SH2 Domains *329*
8.2.2 Phosphotyrosine-binding Domain (PTB Domain) *332*
8.2.3 SH3 Domains *332*
8.2.4 Membrane-targeting Domains: Pleckstrin Homology (PH) Domains and FYVE Domains *334*
8.2.5 Phosphoserine/Threonine-binding Domains *335*
8.2.6 PDZ Domains *336*
8.3 Nonreceptor Tyrosine-specific Protein Kinases *337*
8.3.1 Structure and General Function of Nonreceptor Tyrosine Kinases *337*
8.3.2 Src Tyrosine Kinase and Abl Tyrosine Kinase *338*
8.4 Protein Tyrosine Phosphatases *342*
8.4.1 Structure and Classification of Protein Tyrosine Phosphatases *343*
8.4.2 Cooperation of Protein Tyrosine Phosphatases and Protein Tyrosine Kinases *346*
8.4.3 Regulation of Protein Tyrosine Phosphatases *348*
8.5 Adaptor Molecules of Intracellular Signal Transduction *351*

9 Signal Transmission via Ras Proteins *355*
9.1 The Ras Superfamily of Monomeric GTPases *355*
9.2 General Importance of Ras Protein *358*
9.3 Structure and Biochemical Properties of Ras Protein *360*
9.3.1 Structure of the GTP- and GDP-bound Forms of Ras Protein *361*
9.3.2 GTP Hydrolysis: Mechanism and Stimulation by GAP Proteins *363*
9.3.3 Structure and Biochemical Properties of Transforming Mutants of Ras Protein *366*
9.4 Membrane Localization of Ras Protein *366*
9.5 GTPase-activating Protein (GAP) in Ras Signal Transduction *368*
9.6 Guanine Nucleotide Exchange Factors (GEFs) in Signal Transduction via Ras Proteins *369*
9.6.1 General Function of GEFs *369*
9.6.2 Structure and Activation of GEFs *369*
9.7 Raf Kinase as an Effector of Signal Transduction by Ras Proteins *373*
9.7.1 Structure of Raf Kinase *373*
9.7.2 Interaction of Raf Kinase with Ras Protein *374*
9.7.3 Mechanism of Activation and Regulation of Raf Kinase *374*
9.8 Reception and Transmission of Multiple Signals by Ras Protein *375*

10 Intracellular Signal Transduction: the Protein Cascades of the MAP Kinase Pathways *383*
10.1 Components of MAPK Pathways *385*
10.2 The Major MAPK Pathways of Mammals *388*

10.2.1 The ERK Pathway *388*
10.2.2 The JNK/SAPK, p38 and ERK5 MAPK Pathways *391*

11 Membrane Receptors with Associated Tyrosine Kinase Activity *395*
11.1 Cytokines and Cytokine Receptors *395*
11.2 Structure and Activation of Cytokine Receptors *396*
11.2.1 Activation of Cytoplasmic Tyrosine Kinases *401*
11.2.2 The Jak-Stat Pathway *405*
11.2.2.1 The Janus Kinases *405*
11.2.2.2 The Stat Proteins *406*
11.3 T and B Cell Antigen Receptors *409*
11.3.1 Receptor Structure *410*
11.3.2 Intracellular Signal Molecules of the T and B Cell Antigen Receptors *411*
11.4 Signal Transduction via Integrins *413*

12 Other Receptor Classes *417*
12.1 Receptors with Intrinsic Ser/Thr Kinase Activity: the TGFβ Receptor and the Smad Proteins *417*
12.1.1 TGFβ Receptor *417*
12.1.2 Smad Proteins *418*
12.2 Receptor Regulation by Intramembrane Proteolysis *422*
12.3 Signal Transduction via the Two-Component Pathway *424*

13 Regulation of the Cell Cycle *429*
13.1 Overview of the Cell Cycle *429*
13.1.1 Principles of Cell Cycle Control *429*
13.1.2 Intrinsic Control Mechanisms *431*
13.1.3 External Control Mechanisms *433*
13.1.4 Critical Cell Cycle Events and Cell Cycle Transitions *434*
13.2 Key Elements of the Cell Cycle Apparatus *434*
13.2.1 Cyclin-dependent Protein Kinases, CDKs *435*
13.2.2 Structure of CDKs and Regulation by Phosphorylation *437*
13.2.3 Cyclins *439*
13.2.4 Regulation of Cyclin Concentration *440*
13.2.5 Structural Basis for CDK Activation *442*
13.2.6 Inhibitors of CDKs: the CKIs *445*
13.2.7 Substrates of CDKs *447*
13.2.8 Multiple Regulation of CDKs *449*
13.3 Regulation of the Cell Cycle by Proteolysis *449*
13.3.1 Targeted Proteolysis by the SCF Complex *451*
13.3.2 Proteolysis during Mitosis: the Anaphase-promoting Complex/Cyclosome *452*
13.4 The G_1/S Phase Transition *453*
13.4.1 Function of the D-type Cyclins *454*
13.4.2 Function of pRb in the Cell Cycle *456*
13.5 Cell Cycle Control of DNA Replication *461*

13.6 The G_2/M Transition and Cdc25 Phosphatase *463*
13.7 Summary of Cell Cycle Progression *465*
13.8 The DNA Damage Checkpoints *466*

14 Malfunction of Signaling Pathways and Tumorigenesis: Oncogenes and Tumor Suppressor Genes *469*
14.1 General Aspects of Tumor Formation *469*
14.1.1 Characteristics of Tumor Cells *469*
14.1.2 Genetic Changes in Tumor Cells *471*
14.1.3 Epigenetic Changes in Tumor Cells *472*
14.1.4 Causes of Oncogenic Mutations *473*
14.1.5 DNA Repair, DNA Damage Checkpoints, and Tumor Formation *474*
14.1.6 Cell Division and Tumor Formation *475*
14.2 Cell Division Activity, Errors in Function of Signal Proteins, and Tumor Formation *475*
14.2.1 The Fate of a Cell: Quiescence, Division, or Death *476*
14.3 Definition and General Function of Oncogenes and Tumor Suppressor Genes *477*
14.3.1 Oncogenes and Proto-Oncogenes *478*
14.3.2 Mechanisms of Activation of Proto-Oncogenes *479*
14.3.3 Examples of the Functions of Oncogenes *482*
14.4 Tumor Suppressor Genes: General Functions *487*
14.5 DNA Repair, DNA Integrity and Tumor Suppression *488*
14.6 The Retinoblastoma Protein pRb in Cancer *490*
14.7 The $p16^{INK4a}$ Gene Locus and ARF *493*
14.8 The Tumor Suppressor Protein p53 *494*
14.8.1 Structure and Biochemical Properties of the p53 Protein *495*
14.8.2 Sequence-Specific DNA Binding of p53 *496*
14.8.3 Genes Regulated by p53 *498*
14.8.4 Activation, Regulation and Modulation of the Function of p53 *500*
14.8.5 Overview of p53 Regulation *502*
14.8.6 The MDM2-p53 Network and Cancer *505*
14.9 The Tumor Suppressor APC and Wnt/β-Catenin Signaling *507*

15 Apoptosis *511*
15.1 Basic Functions of Apoptosis *511*
15.2 Overview of Apoptosis *513*
15.3 Caspases: Death by Proteolysis *515*
15.4 The Family of Bcl-2 Proteins: Gatekeepers of Apoptosis *520*
15.5 The Mitochondrial Pathway of Apoptosis *522*
15.6 Death Receptor-triggered Apoptosis *524*
15.6.1 The Fas/CD95 Signaling Pathway *525*
15.6.2 Tumor Necrosis Factor-Receptor 1 and Apoptosis *527*
15.7 Links of Apoptosis and Cellular Signaling Pathways *528*
15.7.1 PI3-Kinase/Akt Kinase and Apoptosis *529*
15.7.2 The Protein p53 and Apoptosis *530*

Index *533*

1
The Regulation of Gene Expression

1.1
Regulation of Gene Expression: How and Where? A Schematic Overview

The transfer of genetic information from the level of the nucleic acid sequence of a gene to the level of the amino acid sequence of a protein or to the nucleotide sequence of RNA is termed gene expression. The entire process of gene expression in eucaryotes includes the following steps:
– *transcription*: formation of a primary transcript, the pre-mRNA
– *conversion of the pre-mRNA into the mature mRNA*: includes processing, splicing, transport from the nucleus to the cytosol
– *translation*: synthesis of the protein on the ribosome.

The expression of genes follows a tissue- and cell-specific pattern, which determines the function and morphology of a cell. In addition, all development and differentiation events are characterized by a variable pattern of gene expression. The regulation of gene expression thus plays a central role in the development and function of an organism. Because of the multitude of individual processes which are involved in gene expression, there are many potential regulatory sites (Fig. 1.1).

Regulation of Transcription

At the level of transcription, it can be determined whether a gene is transcribed at a given point in time. The chromatin structure plays an important role in this decision. Chromatin structures exist that can effectively inhibit transcription and shut down a gene. This "silencing" of genes can be transient or permanent and is generally observed in development and differentiation processes. The regulated transcription of genes requires as an essential step a reorganization and modification of the chromatin, which is a prerequisite for the initiation of transcription and is influenced by epigenetic changes in the DNA in the form of methylation of cytidine residues. Following chromatin reorganization and modification, transcription initiation requires the selection of the target gene and formation of a transcription initiation complex at the starting point of transcription. A large number of proteins are involved in this step. The main components are the multisubunit RNA polymerase, general and specific transcription factors, and cofactors that help to coordinate the chromatin

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

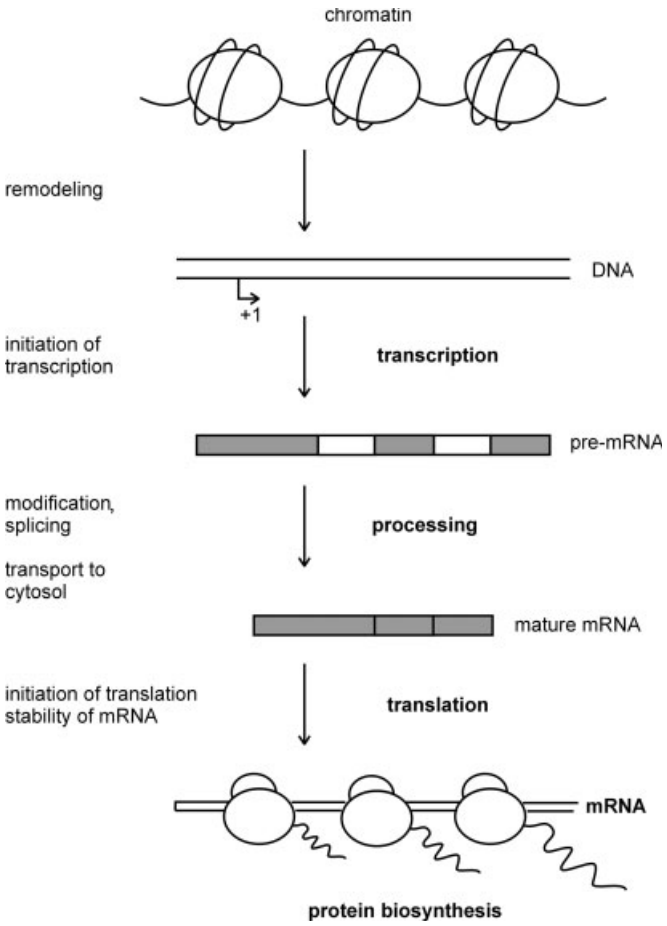


Fig. 1.1 Levels of regulation of eucaryotic gene expression.

structural changes and the process of RNA synthesis. The formation of a functional initiation complex is often the rate-limiting step in transcription and is subject to a variety of regulation mechanisms.

Conversion of the pre-mRNA into the Mature mRNA

Transcription of genes in mammals often initially produces a pre-mRNA, whose information content can be modulated by subsequent polyadenylation or splicing. Various final mRNAs coding for proteins with varying function and localization can be produced in this manner starting from a single primary transcript.

Regulation at the Translation Level

The use of a particular mature mRNA for protein biosynthesis is also highly regulated. The regulation can occur via the accessibility of the mRNA for the ribosome or via the

initiation of protein biosynthesis on the ribosome. In this manner, a given level of mature mRNA can specifically determine when and how much of a protein is synthesized on the ribosome.

Nature of the Regulatory Signals

Regulation always implies that signals are received, processed and translated into a resulting action. The nature of the signals which are employed in the course of the regulation of gene expression and are finally translated into a change in protein concentration can vary dramatically. Regulatory molecules can be small molecular metabolites, hormones, proteins or ions. The signals can be of external origin or can be produced within the cell. External signals originating from other tissues or cells of the organism are transferred across the cell membrane into the interior of the cell, where they are transduced by sequential reactions to the level of transcription or translation. Complex signal chains are often involved in the transduction.

1.2 Protein-Nucleic Acid Interactions as a Basis for Specific Gene Regulation

A recurring motif on the pathway of information transfer from gene to protein is the binding of proteins to DNA or RNA. At the DNA level, specific DNA-binding proteins aid in the identification of genes for regulation via transcriptional activation or inhibition. At the RNA level, specific RNAs are recognized in a sequence-specific manner to attain a controlled transfer of genetic information further on to the mature protein.

The basis of all specific regulation processes at the nucleic acid level is the recognition of nucleotide sequences by binding proteins. For the regulation of gene activity the specific binding of proteins to double-stranded DNA is of central importance. A specific DNA-binding protein usually recognizes a certain DNA sequence, termed the *recognition sequence* or *DNA-binding element*. Because of the enormous complexity of the genome, the specificity of this recognition plays a significant role. The binding protein must be capable of specifically picking out the recognition sequence in a background of a multitude of other sequences and binding to it. The binding protein must be able to discriminate against related sequences which differ from the actual recognition element at only one or more positions.

In the following, the basic features of specific recognition of DNA sequences by DNA-binding proteins will be presented.

1.2.1 Structural Motifs of DNA-binding Proteins

DNA-binding proteins contact their recognition sequences via defined structural elements, termed DNA-binding motifs. DNA-binding motifs are often found in structural elements of the protein which can fold independently from the rest of the protein and therefore represent separate DNA-binding domains.

The region of the binding protein which interacts with the recognition sequence often displays a characteristic small structural element which is stabilized through the help of other structural elements and is thereby brought into a defined position relative to the DNA. These structural elements contain short α-helical or β-sheet structures that in most cases contact the DNA sequence within the major groove: the dimensions of the major groove make it well suited to accept an α-helix. Accordingly, α-helices are often utilized as recognition elements. There are, however, examples of interactions with the minor groove of the double helix (TATA box-binding protein, see Section 1.2.3 and Fig. 1.9). We also know of DNA-binding proteins in which β-structures or flexible structures are involved in contact with the DNA.

The most common and well-characterized DNA-binding motifs can be characterized as described below.

Helix-turn-helix Motif

The helix-turn-helix motif (HTH motif) is – historically seen – the first DNA-binding motif whose structure could be solved in a complex with DNA. It is often found in bacterial repressors. Many eucaryotic DNA-binding proteins also utilize the helix-turn-

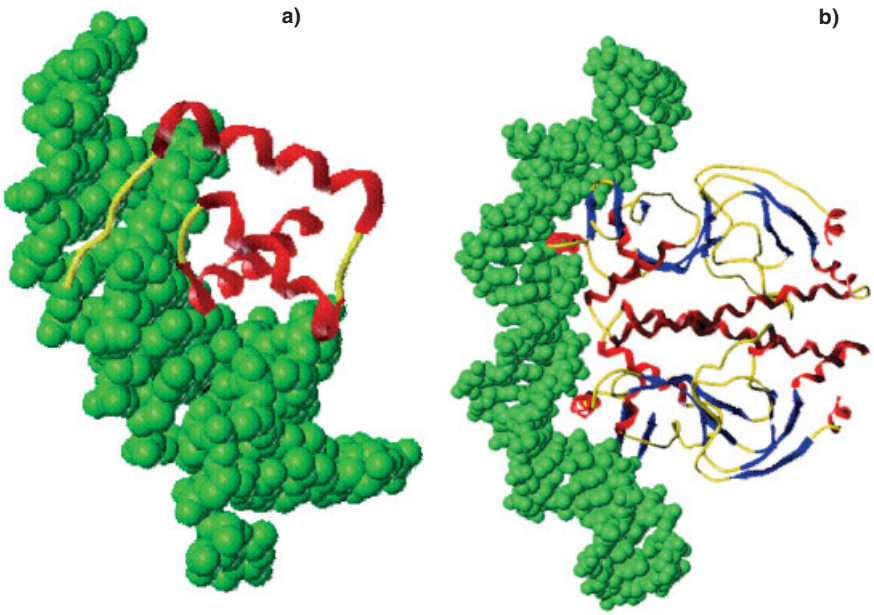


Fig. 1.2 The helix-turn-helix motif in complex with DNA

a) side view of the eucaryotic DNA binding motif of transcription factor MSX-1 in complex with DNA (Hovde et al., 2001). The homeodomain of MSX-1, which is a subgroup of helix-turn-helix motifs, shows a recognition helix (red) embedded in the large groove of the DNA (green). Two additional α-helices stabilize the arrangement of the recognition helix.

b) side view of the complex of a dimer of the catabolite activating protein (CAP) from *E. coli* in complex with DNA. This view displays the symmetrical embedding of two recognition helices, each from one CAP momomer, into the large groove of the DNA element. The DNA is bent nearly 90 deg in the complex. α-helices are in red, turns in yellow and β-strands in blue.

helix motif for specific binding on the DNA. Characteristic of the helix-turn-helix motif is the positioning of an α-helix in the major groove of DNA (Fig. 1.2). The recognition helix is connected by a turn to another helix, whereby the position of the recognition helix is fixed. The two helices occur at an angle of 120° to one other. The binding motif is usually stabilized by further helices of the same or another subunit. The detailed arrangement can differ significantly among the various helix-turn-helix motifs.

Binding Motifs with Zinc Ions

The zinc binding motifs contain Zn^{2+} complexed by four ligating Cys and/or His residues. Based on the stoichiometry of the complex, zinc fingers of the type Zinc-Cys_2His_2, Zinc-Cys_4 and $Zinc_2$-Cys_6 can be distinguished (Fig. 1.3).

The structures of two Zn-binding motifs are shown in Fig. 1.4. The zinc binding motifs play, above all, a structuring role by ensuring that a recognition helix is correctly oriented and stabilized. The zinc ion does not contact the DNA directly.

Fig. 1.5 shows the zinc binding motiv of Zif268, a regulatory DNA-binding protein of mice, in complex with DNA (Pavletich and Pabo, 1991). In Zif268, three zinc-fingers are arranged along the coil of the DNA. The DNA-binding element contains three repeats of the recognition sequence. This results in a modular construction of the protein, so that the periodicity of the DNA is reflected in the protein structure. Another example is found in the DNA-binding domain of the steroid hormone receptors which contains two non-equivalent zinc-Cys_4-motifs (see Section 4.3.2 and Fig. 4.6). In

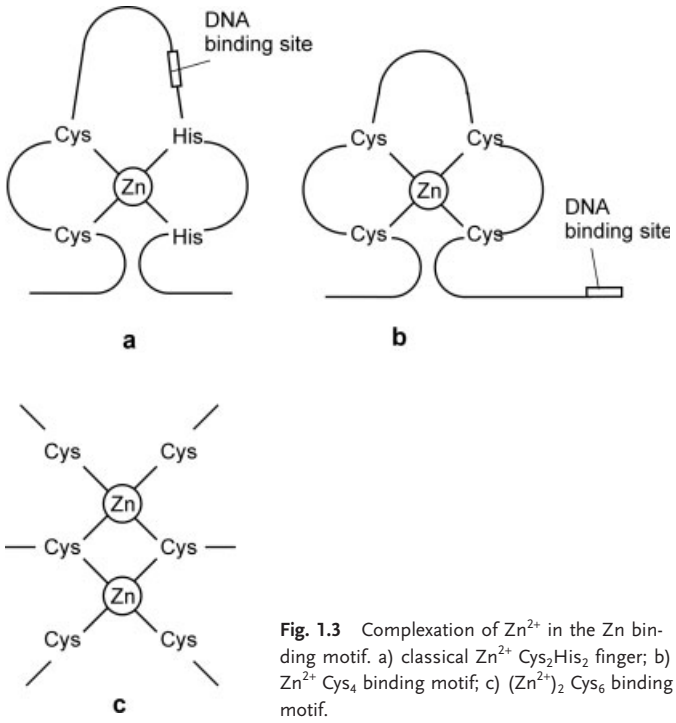


Fig. 1.3 Complexation of Zn^{2+} in the Zn binding motif. a) classical Zn^{2+} Cys_2His_2 finger; b) Zn^{2+} Cys_4 binding motif; c) $(Zn^{2+})_2$ Cys_6 binding motif.

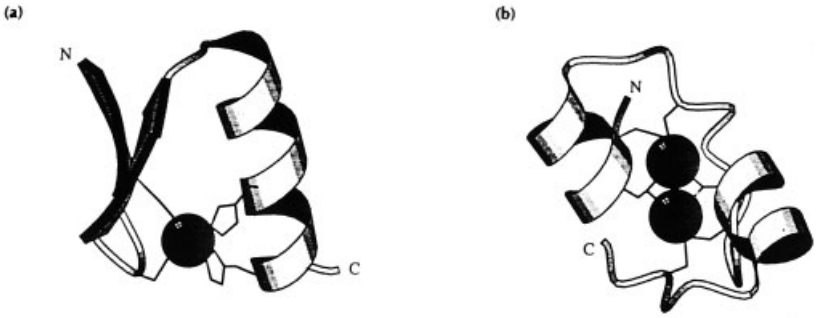


Fig. 1.4 Structures of Zn^{2+} binding motifs. a) TFIIIA-like Zn^{2+} Cys_2His_2 finger; b) the binuclear $(Zn^{2+})_2$ Cys_6 motif of the GAL4 transcription activator.

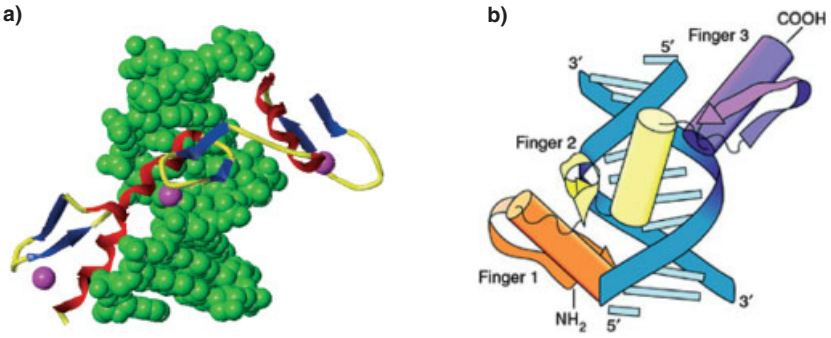


Fig. 1.5 The Zinc binding motif of Zif268 in complex with DNA
a) periodic arrangement of three Zn-fingers in the major groove of the DNA. Two of the Zn^{2+} ions are shown as violet spheres, the recognition helices are in red.
b) schematic drawing of the three Zn-fingers of Zif 268 in complex with DNA.

the transcriptional activator GAL4 of Yeast, two zinc ions are complexed by 6 Cys residues, whereby two of the Cys residues bind to both Zn^{2+} ligands (see Fig. 1.4).

Overall, the zinc-binding motifs display a great variety of structural diversity. The occurrence of a zinc-binding motif can often be predicted based solely on a characteristic series of Cys and His residues in a protein sequence. The complexation of a Zn^{2+} by His and Cys residues serves to bring the recognition element of the protein into a stable and unambiguous position relative to the DNA, thereby enabling specific contacts with the recognition sequence. It has to be pointed out that Zn-binding motifs are also found in many other proteins that do not act on DNA. One example is protein kinase C (see Section 7.4).

Basic Leucine Zipper

This group of binding motifs displays as characteristic structural element an extended bundle of two α-helices that are wound around each other in the form of a "coiled-coil". At their end is a basic region which mediates the DNA binding (review: Ellenberger,

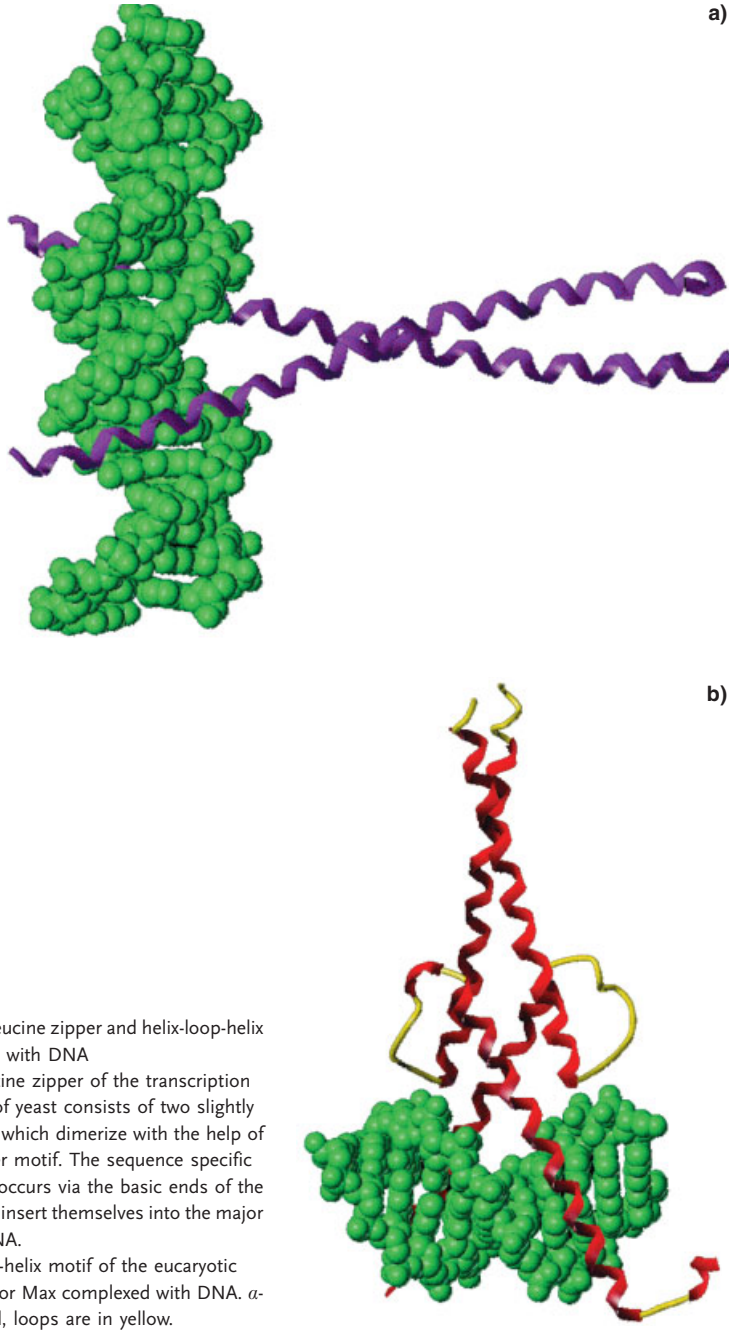


Fig. 1.6 Basic leucine zipper and helix-loop-helix motif in complex with DNA
a) The basic leucine zipper of the transcription activator GCN4 of yeast consists of two slightly curved α-helices, which dimerize with the help of the leucine zipper motif. The sequence specific binding of DNA occurs via the basic ends of the two helices. They insert themselves into the major groove of the DNA.
b) The helix-loop-helix motif of the eucaryotic transcription factor Max complexed with DNA. α-helices are in red, loops are in yellow.

1994). An example of the structure of a basic leucine zipper in complex with DNA is shown by the transcription factor GCN4 from yeast in Fig. 1.6A. *The leucine zipper takes its name from the regular occurrence of leucine residues (or other hydrophobic residues) in an α-helix*. A leucine or other hydrophobic amino acid is found at every seventh position of the helix. This sequential arrangement brings the hydrophobic residues all along one face of the helix, and the hydrophobic residues of two helices can interlock via hydrophobic interaction in a zipper-like manner. The leucine zipper is, above all, a tool to associate proteins in higher dimensions, whereby homodimers as well as heterodimers can be formed. The oligomerization of DNA-binding proteins is usually a prerequisite for strong binding to the cognate DNA element.

The leucine zipper itself does not participate in the recognition; it is only utilized for dimerization of the proteins. The N-terminal end of the basic leucine zipper motif is relatively unstructured in the absence of DNA. A helical structure is induced upon binding to DNA, allowing specific contacts to the recognition sequence. Dimer formation is a prerequisite for the exact positioning of the N-terminal basic end in the major groove of the DNA. Analogous to the dimeric structure of the protein, the DNA sequence displays 2-fold symmetry (see Section 1.2.4).

The Helix-Loop-Helix Motif

One example of the basic helix-loop-helix motif (HLH-motif) is found in the eucaryotic transcription factor Max (Fig. 1.6b and Section 14.3.3). The DNA binding occurs by a parallel bundle of 4 helices with two basic ends. As with the basic leucine zipper motif, the basic ends only attain a defined structure upon binding the DNA. The 4-helix bundle forms via dimerization of two subunits of the Max protein.

DNA Binding via β-Sheet Structures

β-sheet structures as DNA-binding motifs are found in pro- and eucaryotic DNA-binding proteins. Fig. 1.7 shows the structure of the eukaryotic transcription factor NFκB bound to its cognate DNA element. Noteworthy is the enshrouding of the DNA by the β-sheets of NFκB. The recognition of the DNA elements is achieved by interaction of the β-strands with the major groove of the DNA.

Flexible Structures in DNA-Binding proteins

A series of DNA-binding proteins utilize additional flexible structures aside from defined structural DNA-binding motifs in order to increase the stability and specificity of the complex. The λ repressor grabs around the DNA helix with the flexible N-terminal arm of the protein to contact the back of the helix. The basic region of the leucine zipper and HLH-binding proteins is a further example of the importance of protein flexibility in DNA binding. In the absence of DNA, the basic portion of this binding motif is poorly structured, and only following DNA binding is an α-helix formed in the basic region. The α-helix induced upon binding lies in the major groove of the DNA and establishes specific interactions with the recognition sequence.

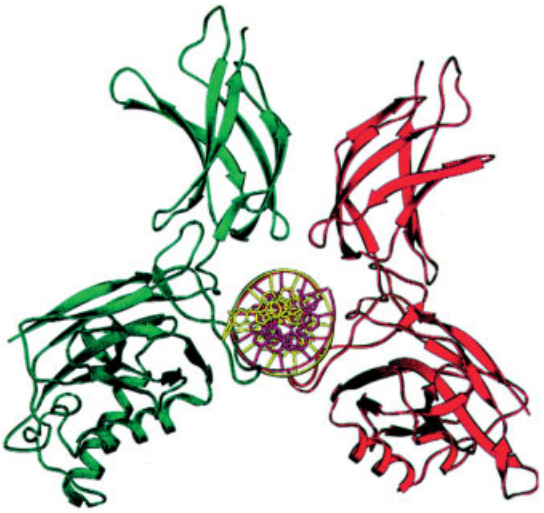


Fig. 1.7 The eucaryotic transcription factor NFκB in complex with DNA. Shown is the structure of the p50–p65 heterodimer of NFκB complexed with a specific NFκB DNA element. The view is along the DNA helical axis. Each of the subunits contains a bundle of β-sheets which envelops the DNA so that only the minor groove is exposed. The p50 subunit is shown in green and the p65 subunit in red The top strand of DNA is shown in pink, the bottom strand in yellow. From Chen et al., 1998; used with permission.

1.2.2
The Nature of the Specific Interactions in Protein-Nucleic Acid Complexes

The binding of a protein to nucleic acid is accomplished by weak, noncovalent interactions. The interactions are the same as those involved in the formation of the tertiary structure of a protein:
- hydrogen bonds (H-bonds)
- electrostatic interactions
- van der Waals interactions
- hydrophobic interactions.

H-bonds in Protein-Nucleic Acid Complexes

Of central importance for the formation of a specific protein-DNA complex are hydrogen bonds. The H-bonds are clearly identifiable in high-resolution structures. H-bonds occur where an H-bond donor and acceptor lie within 0.27–0.31 nm of each other. Energetically most favorable is the linear arrangement of the H-bond, with deviations from linearity leading to a reduction in energy. This characteristic is responsible for the stereospecific orientation of H-bond acceptors and donors. The H-bond thus contributes significantly to the spatial orientation between protein and nucleic acid.

There are many different H-bond donors as well as acceptors in proteins and nucleic acids which contribute to the specific recognition. Important H-bond donors and acceptors in proteins are Asn, Gln, Ser, Thr, Tyr, Glu, Asp, Arg, Lys, Cys and His. The peptide bonds of the backbone often participate also.

The heteroatoms and exocyclic functional groups of the bases within the nucleic acid can form H-bonds to residues of a binding protein, in addition to base pairing. Also, the oxygen of the ribose or deoxyribose and the phosphate moiety of DNA can be used as H-bond acceptors.

The available structural information on protein-DNA complexes reveals great variability and flexibility in the spectrum of H-bond interactions. Examples of the variety of H-bond interactions are shown in Fig. 1.8.

The following points are noteworthy:

- A base can be contacted by more than one amino acid residue. Furthermore, there are many examples of one amino acid residue, e.g. Arg, contacting two sequential bases. This type of interaction functions as a clip and maintains a spatially defined arrangement.
- The contact between protein and DNA can also be transmitted via bound water molecules.
- There are always numerous H-bond contacts formed between the recognition sequence and the binding protein. The pattern of H-bond donors and H-bond acceptors is determined by the sequence and conformation of the DNA as well as by the specific structure of the protein. Both together lay the foundation for a specific recognition of the DNA by the protein.
- An important factor in the structure of protein-DNA complexes can be the peptide backbone. The amide bond can function as an H-bond acceptor as well an H-bond donor. Because of the reduced flexibility of the backbone vs side chain (resonance stabilization of the peptide bond), H-bonds to the peptide backbone lead to a rigid and tight arrangement in the complex and contribute extensively to the exact fit between protein and nucleic acid.

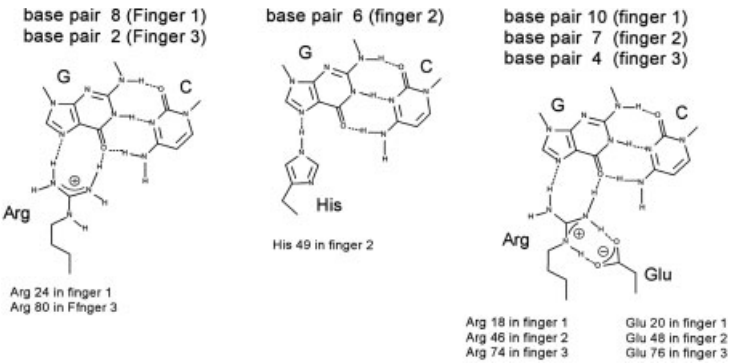


Fig. 1.8 H-bonds in the complex between the Zinc fingers of Zif268 with the cognate recognition helix. Zif268 contacts the DNA with three Zn-fingers (finger 1-3 in Fig. 1.5). Shown are some H-bond contacts formed between the fingers and the base pairs of the recognition sequence.

Ionic Interactions

Ionic interactions result from the electrostatic attraction or repulsion between charged groups. As opposed to H-bonds, ionic interactions are not directed and are effective over greater distances.

DNA presents itself to a binding protein as a negatively charged, anionic substrate. Accordingly, the protein displays a complementary positive potential resulting from an accumulation of basic amino acid residues. The electrostatic interaction between the two oppositely charged binding surfaces of DNA and protein make a significant energetic contribution to the formation of a stable complex.

The ionic interactions are, however, less suitable for distinguishing between various base pairs, since only the phosphates of the backbone from the DNA are involved in the interaction. Together with the specific H-bonds, the nonspecific ionic interactions contribute significantly to the formation of a stable complex. The positively charged surface of DNA-binding proteins is also the reason for the ability of many such proteins to bind DNA nonspecifically.

Van der Waals Contacts

The van der Waals contacts are a type of electrostatic interaction and arise from an interaction between permanent and/or induced dipoles in the bond pair. They are typically effective over a much shorter range than ionic interactions. The contribution of van der Waals contacts to the binding of a protein to a DNA sequence is difficult to estimate, since many small contributions must be considered. An example of a contact surface with many van der Waals interactions can be found in the complex of the TATA box-binding protein with the TATA box (see Fig. 1.9). In this case there are extensive van der Waals contacts between the sugar residues of the DNA backbone and the hydrophobic surface of the protein (Kim et al., 1993).

1.2.3 The Role of the DNA Conformation in Protein-DNA Interactions

The double helix of the DNA can only to a first approximation be considered a linear, rod-like structure with the typical coordinates of B-DNA. Actually, DNA possesses considerable flexibility and conformational variability. The flexibility and structural polymorphism of DNA are prerequisites for many of the regulatory processes on the DNA level (review: Alleman and Egli, 1997). Local deviation from the classical B-structure of DNA, as well as bending of the DNA, are observed in most protein-DNA complexes.

Local Conformational Changes of DNA

Crystal structures of DNA have shown that, apart from the structural motifs of the A-, B- and Z-forms of DNA, other, sequence-dependent structural variations exist which are observed when smaller sequence fragments are examined in detail. The structural variations can affect the width of the major groove, the extent of base stacking, and the tilt of the base pairs to each other. The local conformational changes are sequence

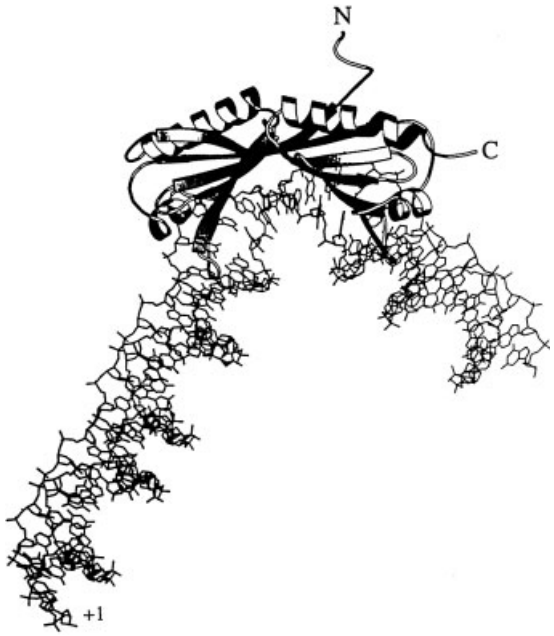


Fig. 1.9 Bending of DNA in the TATA box. The DNA is kinked in the complex of the TATA box binding protein (yeast) with the 8 base pair TATA box (Kim et al., 1993). The DNA is deformed in the region near the kink: the minor groove, which faces the protein, is clearly widened.

dependent and can be intrinsic properties and thus permanent occurrences; they can, however, also be induced by protein binding.

In most protein-DNA complexes, analysis of DNA structure in the region of contact with the binding protein reveals distinct divergence from the parameters of classical B-DNA structure. A specific sequence-determined conformation of the DNA is often a prerequisite for a specific recognition.

Bending of DNA

If one traces a longer stretch of a DNA molecule in solution, a clear divergence from linearity becomes evident. Thermally induced structural fluctuations allow a bending of DNA, which is why long DNA molecules are described as a random coil. This bending of the DNA occurs in molecules with a length of more than ca. 200 bp.

Bending of shorter fragments is observed in the presence of distinct sequence characteristics or upon binding of proteins. An intrinsic bending of short DNA fragments is induced when the DNA contains short dA-repeats (e.g. dA_5), and this bending can be enforced by protein binding.

Protein-induced Bending of DNA

There are numerous examples of protein-induced bending of DNA. The bending of a short segment of DNA (150-200 bp) leads to a loss of stacking interactions of the π-electron system of neighboring bases and is energetically unfavorable. Stacking interactions arise from interactions of the π-electron systems of bases atop one another and contribute extensively to the stability of the double helix. An active bending of a short

piece of DNA is therefore only possible if the energy loss is compensated for by other favorable interactions. For protein-induced bending of DNA, the energy is provided by the complex formation with the protein. A portion of the favorable interaction energy (H-bonds, hydrophobic interaction, ionic interactions) compensates for the energy required to bend the DNA.

The divergence of the DNA conformation from a rod-like structure is observed to a variable extent. The DNA can be slightly curved, as observed in nucleosome-bound DNA, or it can be abruptly kinked as shown in Fig. 1.9 for the TATA box-binding protein.

The TATA box-binding protein causes a kinking of the bound DNA at an angle of ca. 100° (Fig. 1.9). The flexibility of the alternating purine-pyrimidine sequences of the binding site favor a prominent deformation of the DNA with little energy requirement. Thus, in the region of the kink, the minor groove is widened and the DNA strands partially separated. The widening of the minor groove allows numerous van der Waals contacts with the protein.

Regulatory processes at the protein-DNA level generally occur in multiprotein complexes where several proteins interact with distinct DNA elements and cooperate over a larger distance of DNA. This requires, above all, communication between various DNA-bound proteins, which may not be bound to neighboring sequences. An important role of the actively induced or intrinsic DNA bending is to bring together linearly separated DNA sequences and hence bring together their bound protein (Fig. 1.10). Only by bending DNA does an effective interaction between DNA-binding proteins bound to distant DNA-binding elements become possible. As an example, a highly ordered DNA multi-protein complex is formed during transcription initiation. A crucial role is ascribed to the TATA box-binding protein in the assembly of the initiation complex by inducing bending of the bound DNA. This bending creates defined binding sites for other components of the transcription initiation complex. Furthermore, other sites, which are separated in their linear sequence, are brought into close proximity. The bending of DNA therefore plays an essential role in gene activation.

1.2.4
Structure of the Recognition Sequence and Quaternary Structure of DNA-binding Proteins

The recognition sequences for specific DNA-binding proteins usually include only 3–8 base pairs, arranged either palindromically or in direct repeats (Fig. 1.11). The symmetry of the sequence in the DNA element is often reflected in the subunit structure of the binding protein. Less common is the occurrence of a singular recognition sequence.

Palindromic Arrangement

Palindromic sequences with 2-fold symmetry are usually bound by dimeric proteins in which each subunit of the protein contacts one half-site of the DNA element. The use of twofold symmetry in the binding sequence and the protein dimers is an economical approach to achieving high-affinity binding. The DNA-binding motif of one subunit

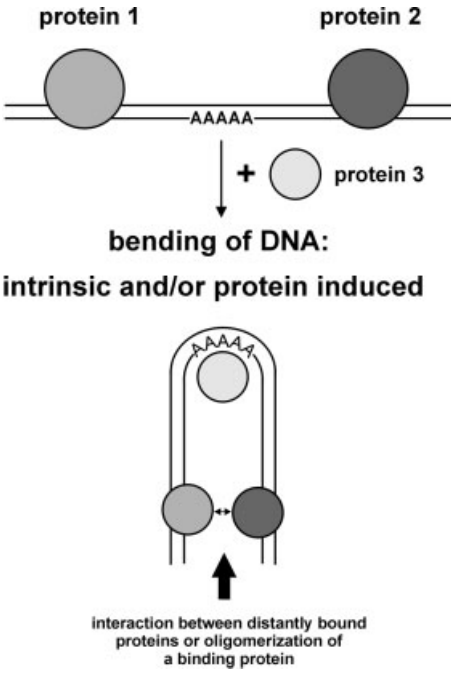


Fig. 1.10 The significance of the bending of DNA for protein-protein interactions. DNA-bound proteins, which would not interact if associated with linear DNA, can be brought together through intrinsic or protein-induced bending of the intervening sequences. The bending of the DNA creates a high local concentration of the two proteins and thus enables their effective interaction.

often contacts only a few base pairs of the recognition sequence when in a complex. This is generally not sufficient to ensure tight binding of a subunit. Because of the repetition of the recognition sequence in a DNA element, binding by the two subunits of a dimeric binding protein occurs in a *cooperative manner*. If one subunit of a protein contacts one half of the recognition sequence, then binding by the other subunit to the other half is strongly favored. Both subunits bind cooperatively, and a high-affinity binding results. The twofold symmetry in the DNA sequence and binding protein plays an important role in the specific binding process. If, for example, a mutation inactivates one half of the recognition sequence, the other intact site often no longer suffices to provide for a tight binding. The protein can then only bind weakly, and the mutated DNA element is often inactive in the *in vivo* situation.

Direct Repeats of the Recognition Sequence

Direct repeat of the recognition sequence requires a nonsymmetrical spatial arrangement of the bound protein subunits (see Chapter 4: Nuclear Receptors). The protein-DNA complex has, in this case, a polar character, and the two proteins bound on the two respective halves of the DNA element can carry out different functions. Direct twofold repeats are observed for the DNA-binding elements of members of the nuclear receptor superfamily (see Chapter 4).

The promotor regions of eucaryotes often contain multiple repeats of DNA elements. In this case, there can be a tandem-like arrangement of the oligomers of a DNA-binding protein, allowing cooperative interactions and formation of higher-order complexes.

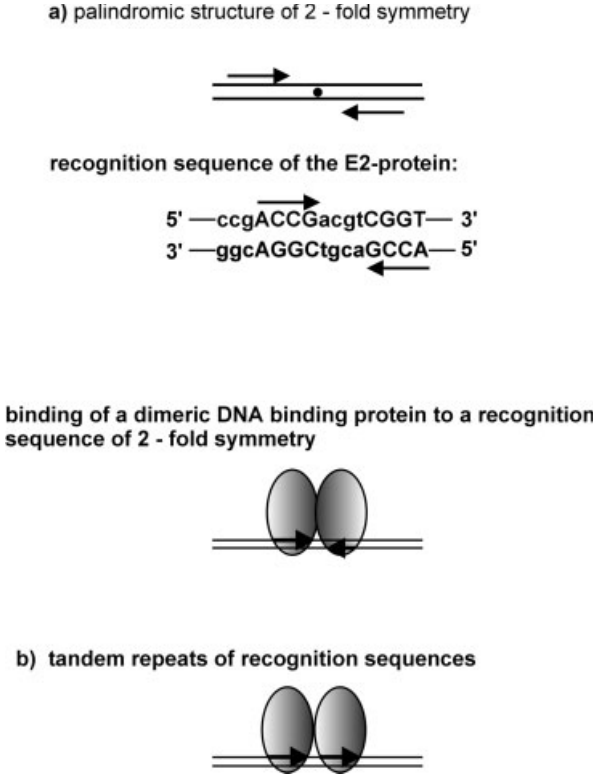

a) palindromic structure of 2 - fold symmetry
recognition sequence of the E2-protein:
5' —ccgACCGacgtCGGT— 3'
3' —ggcAGGCtgcaGCCA— 5'
binding of a dimeric DNA binding protein to a recognition sequence of 2 - fold symmetry
b) tandem repeats of recognition sequences

Fig. 1.11 Structure and symmetry of DNA recognition elements and the oligomeric structure of DNA binding proteins.The symmetry of the DNA sequence and the binding protein plays an important role in the specific binding process. If, for example, a mutation inactivates one half of the recognition sequence, the other intact site often no longer suffices to provide for a tight binding. The protein can then only bind weakly and the mutated DNA element is often inactive in the ***in vivo*** situation.

The occurrence of tandem-like repeats of the DNA elements, in conjunction with the oligomerization of the cognate DNA-binding protein, allow specific structures to be created which are vital for further regulatory processes. This functional principle is demonstrated poignantly by the *E. coli* Lac repressor/operator system. Three Lac repressor-binding sites are found within a 500 bp stretch in the Lac operon of *E. coli*. Each of the three binding sites has a two-fold palindromic structure on which the Lac repressor binds as a dimer (Lewis et al., 1996). The Lac repressor, however, exercises its full repressive function as a tetramer. It is therefore assumed that dimers bound to adjacent binding sites associate into tetramers (Fig. 1.12). The intervening sequence from 93 to 401 forms a loop, also termed the "repression loop". The repressor acts in this arrangement as a clip to bring together widely separated DNA sequences. It is assumed that the specific arrangement of DNA in the loop has decisive consequences for the ensuing transcription activity: the binding of RNA polymerase is hindered by

the DNA loop, while the loop creates the structural framework for further regulatory proteins to bind, e.g., the CAP protein.

The distance between and nature of the bases in both the palindromic and the direct repeats of the recognition sequence plays an important role. It is evident that a dimeric protein would bind optimally to a twofold symmetric sequence only if the distance between the recognition elements matches the distance as determined by the protein structure. If one increases the distance by a few base pairs, a loss in cooperative binding capacity of the dimerization motifs of the protein to the rigid intervening DNA may

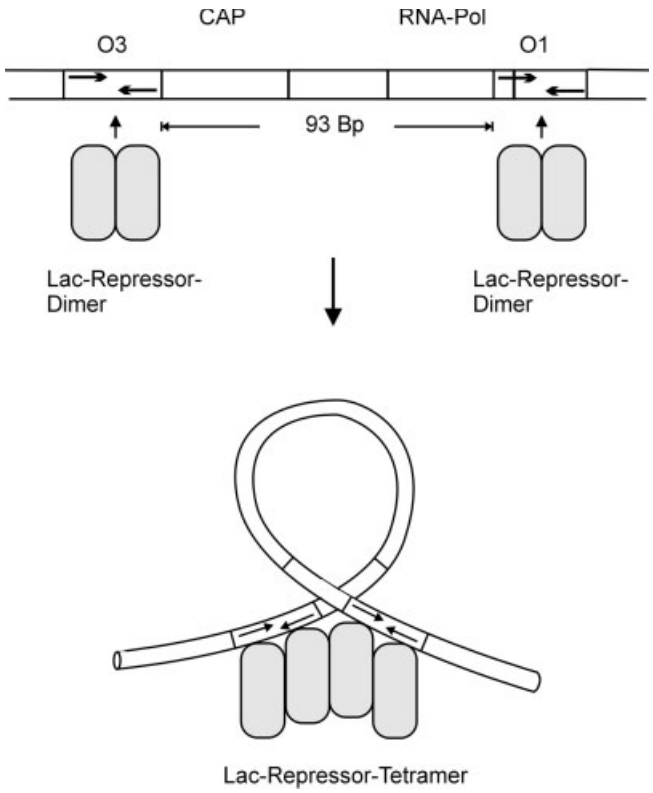


Fig. 1.12 Tetramerization of the Lac repressor and loop formation of the DNA. The Lac repressor from *E. coli* binds as a dimer to the two-fold symmetric operator sequence, whereby each of the monomers contacts a half-site of a recognition sequence. The Lac operon of *E. coli* possesses three operator sequences O1, O2 and O3, all three of which are required for complete repression. O1 and O3 are separated by 93 bp, and only these two sequences are displayed in the figure above. Between O1 and O3 is a binding site for the CAP protein and the contact surface for the RNA polymerase. The Lac repressor acts as a tetramer. It is therefore assumed that two dimers of the repressor associate to form the active tetramer, whereby one of the two dimers is bound to O3, the other dimer binds to O1. The intervening DNA forms a so-called repression loop. After Lewis et al., 1996.

result. The distance between the contacting sequence elements is particularly important for the DNA-binding elements of the nuclear receptors (see Chapter 4). The DNA-binding element of the estrogen receptor differs from, e.g., that of the T_3-receptor only with respect to the number of bases between the two half-sites of the recognition sequence. In this case the distance between the half-sites decides which of the two receptors will bind and act as gene regulators.

A further aspect of the occurrence of multimeric recognition elements is the possibility of the formation of *heterodimers* (see Section 1.4.5.3). There exist related classes of DNA-binding proteins which recognize similar DNA-binding motifs and possess a common dimerization motif. Among these, both homodimers and heterodimers can be formed, which bind to DNA with slightly different specificities. The possibility of the formation of heterodimers and homodimers of related DNA-binding proteins represents an important strategy for expanding the specificity of the regulatory process. A notable example is the nuclear receptors (Chapter 4).

1.3 The Principles of Transcription Regulation

1.3.1 Elements of Transcription Regulation

Transcription represents the most important point of attack for the regulatory processes which control the flow of genetic information from DNA to mature protein. Primarily, it is the initiation of transcription that is regulated, since this represents the rate-limiting step. The essential elements of regulation at the level of initiation in eucaryotes are (Fig. 1.13)

- *cis*-acting DNA-sequences
- trans-acting DNA-binding proteins
- structure of chromatin.

Cis-acting DNA sequences usually represent specific protein-binding sites that lie near the start site of transcription or are quite distanced from it. Furthermore, there are examples among eucaryotes in which the *cis* element is found within the transcribed region. Protein binding to the *cis*-acting elements can have an activating or an inhibitory effect on transcription. If the activating *cis*-element is located far from the site of action and its effect is also orientation-independent, then it is termed an enhancer. Inhibitory *cis*-elements of this type are called silencers. Furthermore, one frequently observes in eucaryotes so-called *composite control regions*, which contain various *cis* elements. In this case, several transcription factors act cooperatively in the initiation of transcription.

Trans-acting DNA-binding proteins specifically bind the *cis*-elements DNA to thereby select the gene to be transcribed. These proteins can exercise a negative or positive influence on transcription upon binding to their cognate DNA sequence.

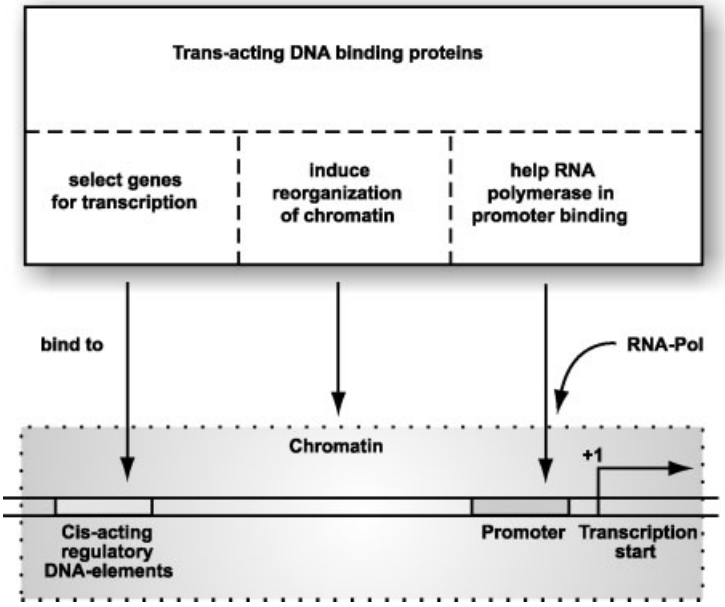


Fig. 1.13 The main functions of trans-acting DNA-binding proteins in transcriptional regulation

Chromatin structure is a major attack point for transcription regulation in eucaryotes. Efficient transcription initiation requires a specific structure and modification of chromatin, which can be positively or negatively controlled by the *trans*-acting proteins.

Negative Regulation of Transcription

Negative regulation of transcription implies that the binding of a regulatory protein leads to inhibition of transcription. Such proteins are described as transcriptional *repressors*. Negative regulation among procaryotes is often accomplished by the bound repressor blocking the access of the RNA polymerase to the promotor. This occurs if, for example, the binding sequence of the repressor and promotor sequence partially overlap. Bound repressor proteins can also cause a change in the conformation and topology of the DNA, which can indirectly inhibit transcription. Another mechanism for negative control involves binding of the regulatory protein to another protein whose function is essential for transcription; such binding then interferes with the function of the latter (see Section 1.3.3.2 and 1.4.5). Furthermore, specific chromatin structures can exert a repressive effect on transcription (see Section 1.4.7).

Positive Regulation of Transcription

Positive regulation implies that the bound protein stimulates transcription. Such proteins are termed transcriptional activators. There are various mechanisms of transcriptional activation. Usually protein-protein interactions between the transcriptional activator and components of the transcription apparatus or chromatin are in-

volved. Activation of transcription can also be achieved via changes in chromatin structure.

1.3.2 Functional Requirements for Repressors and Transcriptional Activators

Regulatory DNA-binding proteins are multi-functional. Aside from their DNA-binding property, they also have the ability to register regulatory signals and transmit these on to the transcription apparatus (Fig. 1.14).

Specific DNA Binding

Regulatory DNA-binding proteins generally display specific and selective DNA-binding capacity. In this way, only those genes which possess a copy of a particular DNA-binding element are subjected to regulation by the corresponding binding protein.

Registering a Regulatory Signal: Activation and Inactivation

A regulatory DNA-binding protein possesses structural elements for the registration of incoming signal, which leads to a change in concentration of the *active* binding protein. The activation (or inactivation) of the binding protein can be connected with a change in the ability to bind DNA, or can influence the capacity of the protein to interact with the transcription apparatus and with chromatin-modifying proteins.

Communication with the Transcription Apparatus

The DNA-binding protein must be capable of transmitting signals to the transcription apparatus via protein-protein interactions. Distinct regions of transcription factors contain interaction motifs that bind to and recruit protein components of the transcription apparatus. DNA binding alone can be ascribed the function of increasing the effective concentration of the transcription regulator at the site of the transcription apparatus.

Communication with Chromatin

Changes in chromatin structure are at the heart of transcription regulation in eucaryotes. Regulatory DNA-binding proteins communicate with chromatin-modifying and chromatin-remodeling protein complexes to generate specific chromatin structures at promotor regions.

Turning off the Transduction of Signal

Regulatory signals should only be effective for a limited period of time and under certain external conditions. This also holds, of course, for regulation at the transcription level. It is therefore necessary to turn off the transduction of signal by the DNA-binding protein after the mediated demands have been fulfilled. Cells use common mechanisms for both the activation and inactivation of signal pathways. These are summarized below.

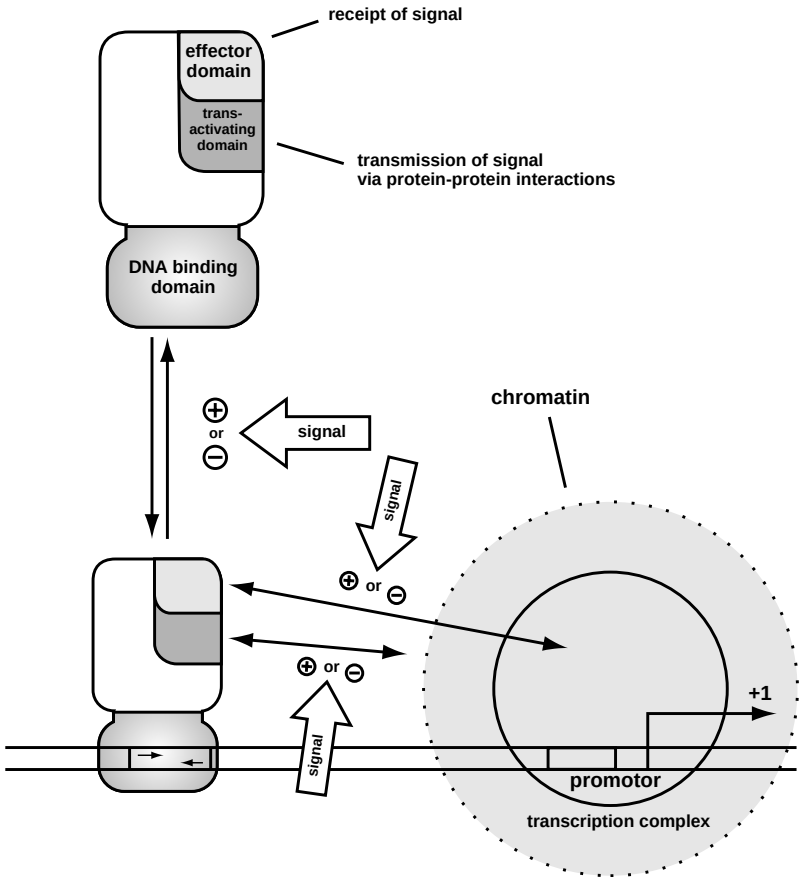


Fig. 1.14 Structural and functional principles of transcription activators. Typical transcription activators of eucaryotes possess a DNA binding domain, an effector domain and a transactivating domain. An incoming signal is registered by the effector domain and transduced to the chromatin and the transcription apparatus. In the active state, the transcription activator induces a chromatin structure that is favorable for transcription. Protein-protein interactions with the transcription apparatus bound to the promotor mediate a stimulation of transcription initiation as well.

1.3.3
Mechanisms for the Control of the Activity of DNA-binding Proteins

Regulatory DNA-binding proteins are controlled by a multitude of mechanisms. These controls operate at the level of the concentration of the binding protein or they act on preexisting DNA-binding proteins by post-translational mechanisms.

In the latter case the control may influence the DNA-binding activity of the protein or it may change the ability of the protein to communicate with the transcription apparatus or with chromatin components.

1.3.3.1 Binding of Effector Molecules

Low-molecular-weight effectors are commonly employed in bacteria to change the DNA-binding activity of repressors or transcriptional activators and to control the amount of active DNA-binding proteins. This type of regulatory mechanism is frequently used for metabolic pathways, as in, for example, the biosynthesis and degradation of amino acids. The effector molecules represent components arising from the particular metabolic pathway. The goal of this regulation is to adjust the transcription rate to the current demand of the gene product.

The binding of low-molecular-weight effectors to regulatory DNA-binding protein can lead to an increase or decrease in the affinity of the protein for its recognition sequence.

The strategies and mechanisms of action of effector molecules on regulatory DNA-binding proteins can be elucidated using the example of the Trp repressor of *E. coli*. In this system binding of the effector increases the affinity of the binding protein to its DNA element.

The Trp repressor controls the transcription of a total of five enzymes required for the biosynthesis of tryptophan (Fig. 1.15a). The genes for the five enzymes are encoded in a single operon, whereby the binding site for the Trp repressor overlaps with the promotor. The bound repressor blocks the RNA polymerase's access to the promotor, thereby inhibiting transcription.

The enzymes of Trp-biosynthesis are only required if too little tryptophan is available to the bacteria from the growth medium. In such a case the Trp requirement is fulfilled by the cell's own Trp biosynthesis. If, however, there is enough Trp supplied by the medium, then it is prudent to shut down the Trp operon. The sensor is the Trp concentration. The Trp repressor registers the current Trp concentration with the help of its own Trp-binding site. If a great deal of Trp is present, then the Trp-binding site of the repressor is occupied by Trp. The Trp repressor binds Trp with high affinity ($K_D = 10^{-9} - 10^{-10}$ M), upon which transcription of the operon is then blocked.

At low Trp concentration, the Trp repressor is mainly in the unbound, inactive form. The free form of the Trp repressor binds with a ca. 10^4-fold lower affinity to the recognition sequence than that of the Trp-bound form. The promotor remains free under these conditions, and transcription of the genes for Trp biosynthesis can occur. The shutting on and off of the Trp operon is based on the disparate DNA affinities of the free and Trp-bound repressor.

The Trp repressor is representative of many other DNA-binding proteins which occur in a binding and nonbinding form as regulated by effector molecules. The binding of the effector molecule determines whether the protein is in the active or the inactive form. Active and inactive forms often differ by a factor of $10^4 - 10^5$ in their affinity for their cognate sequence. The affinity of the inactive form for the recognition sequence usually lies in the same range as its affinity for a random, nonspecific DNA sequence. The inactive binding protein is incapable of selectively binding the specific DNA element. The structural basis for the affinity differences are changes in the protein structure induced upon binding the effector molecule.

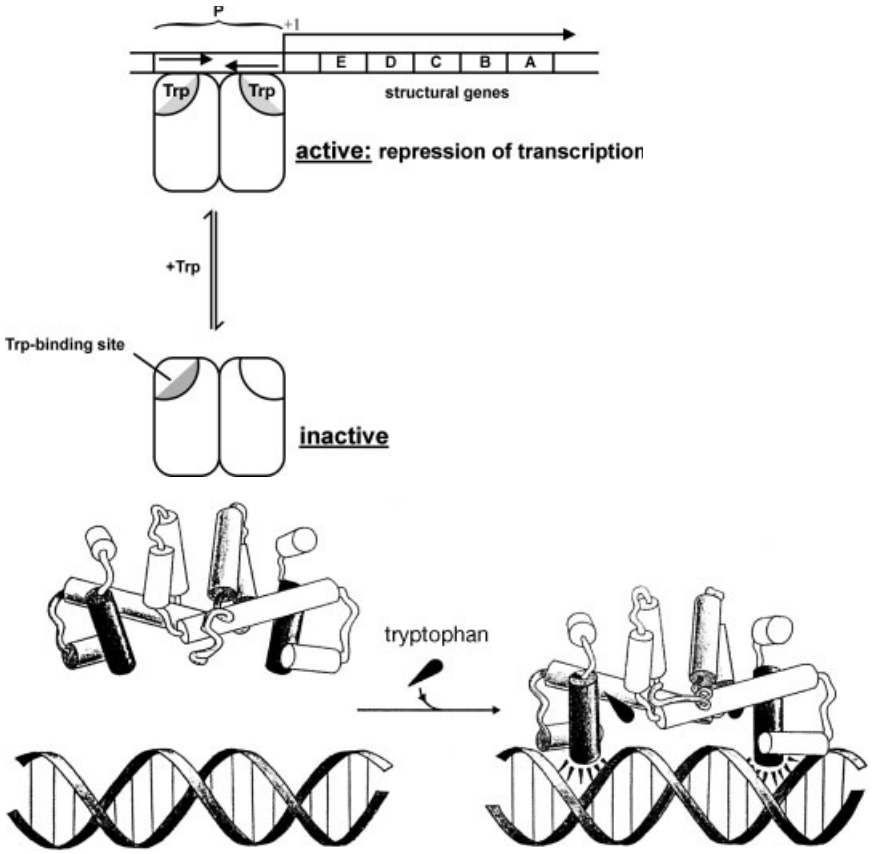


Fig. 1.15 Regulation of the Trp operon in *E. coli*. A) The Trp repressor requires Trp in order to bind its affiliated DNA binding element. In the absence of tryptophan, the Trp repressor can not bind to the regulatory sequence and is therefor inactive. Upon an increase in the tryptohan concentration, tryptophan binds to the Trp repressor and transforms it into a binding-proficient form. The DNA bound Trp repressor prevents the transcription of the structural genes, and the biosynthesis of tryptophan is halted. B) structural basis for the activation of the DNA binding of the Trp repressor by tryptohan molecules. The dimeric Trp-repressor recognizes the two-fold symmetric recognition sequence with the help of a helix-turn-helix motif. In the absence of tryptophan the spacing between the two recognition helices is too small to allow entry into and binding of the major groove. Upon binding tryptophan, the flexible recognition helices are pushed apart and oriented for optimal contact to the half sites of the DNA element. Reproduction with permission from Alberts et al., 1994, p. 418.

Molecular Basis for the Control of Binding Activity of a Repressor by Effector Molecules
Comparison of the structures of a binding protein in the inactive and active forms bound to DNA gives an impression of the conformational changes correlated with binding of effector molecules. The Trp repressor is an example in which the difference between the DNA-binding affinities of the inactive and active forms can be explained in structural terms (Fig. 1.15b).

In the Trp-bound, binding-competent form, the helix-turn-helix motif of the repressor is found in a position favorable for contacting the recognition sequence, and the recognition helix can interact with the major groove of the DNA. The effector molecule tryptophan binds near the helix-turn-helix motif and performs several tasks. First, it orients and fixes the recognition helix in such a way that the specific interactions with the DNA recognition element can be formed. Furthermore, the bound tryptophan is indirectly involved in interactions with the DNA, in that it supports the formation of H-bonds to the DNA by certain amino acid residues. In the Trp-free form, the prerequisite for a strong cooperative binding of the repressor dimer is not fulfilled, since the recognition helices are not positioned optimally for binding to the recognition sequence in the major groove of the DNA.

Metal Ions as Effector Molecules

Metal ions can serve as effector molecules as well as controlling the DNA-binding activity of regulatory proteins. A recent example of the regulation of DNA-binding proteins by metal ions is the transcriptional repressor DREAM, which binds to the cognate DNA element only in the absence of Ca^{2+} (Carrion et al., 1999). An increase of Ca^{2+} in the form of a Ca^{2+} signal (see Chapter 6) leads to a reduced affinity to its DNA element and to an increased expression of the target gene.

1.3.3.2 Binding of Inhibitory Proteins

Specific DNA-binding proteins can be constrained in their ability to function as gene regulators by complex formation with inhibitor proteins. Examples are the steroid hormone receptors, which, in the cytosol, are bound in their inactive form to the proteins hsp90, hsp56 and p23 (see Chapter 4). In response to an incoming signal in the form of increased concentration of steroid hormones, the inhibitory complex is dissolved. The binding of steroid hormone to the receptor enables the dissociation of the inhibitory protein and the subsequent transport into the nucleus, where the receptor can function as a gene regulator (see Section 4.6). Other members of the nuclear receptor superfamily are kept in an inactive, DNA-bound state by corepressors (see Section 4.4). Transition to the active state is achieved by binding of hormone and dissociation of the corepressor.

Protein phosphorylation also serves as a tool to release the DNA-binding protein from an inhibitory complex in the cytosol, as illustrated by the transcription factor NFκB, which is kept in an inactive state by complexation wth the inhibitor IκB. (see Section 2.6.5). Here, incoming signals induce phosphorylation of the inhibitor IκB, leading to its proteolytic destruction and liberating NFκB for transport into the nucleus.

1.3.3.3 Modification of Regulatory Proteins

Post-translational covalent modification of DNA-binding proteins is a mechanism commonly employed among eucaryotes to control the activity of DNA-binding proteins. The following mechanisms stand out:

- **Phosphorylation**
  Phosphorylation at Ser/Thr residues is of particular importance for the regulation of eucaryotic transcription factors. Functional and mechanistic consequences of the phosphorylation of transcription factors will be discussed in more detail in the section on the regulation of eucaryotic transcription (see Section 1.4.5.2). Specific or nonspecific protein phosphatases (see Section 7.6) can remove the phosphate residues and terminate the phosphorylation signal.
- **Methylation, acetylation**
  Covalent modification by methylation or acetylation at Arg or Lys residues can be used to regulate the interaction of transcription factors with other regulatory proteins. As an example, Arg methylation of the transcription factor Stat1 regulates its dephosphorylation by protein tyrosine phosphatases (Zhu et al., 2002). Acetylation of Lys residues controls the activity of the yeast transcription factor GATA-1 (Boyes et al., 1998).
- **Redox-regulation**
  The reversible oxidation of cysteine residues has been shown to function as a switch between different states of activity of transcription factors. This has been shown, e.g., for the transcription factor AP1, which contains cysteine motifs that regulate activity in response to oxidative stress (Karimpour et al., 2002).

1.3.3.4 Changes in the Concentration of Regulatory DNA-binding Proteins

The amount of available DNA-binding proteins is, in many situations, a critical factor for the extent of transcription regulation. The concentration of regulatory DNA-binding proteins can be regulated within the framework of the following processes in eucaryotes:

– transcription
– splicing, transport
– translation
– compartmentalization
– targeted degradation.

The above points will be discussed in more detail in the following section (Section 1.4) in the context of eucaryotic gene regulation.

Only *autoregulation* will be introduced as an example of the regulation of DNA-binding proteins at the level of transcription.

Autoregulation implies that a repressor regulates the transcription of its own gene (Fig. 1.16). In the operator region for the genes of the repressor is found a binding site for the repressor itself, so that it can function as its own negative regulator. If little repressor is available, then the associated DNA element remains unoccupied, and the transcription of the repressor gene is no longer blocked. Increasing concentration of the repressor leads to increased occupation of the repressor-binding site and to an inhibition of the transcription of the repressor. Usually, binding sites also exist for the repressor in other operons. The extent of occupation of the various operons is determined by the affinity of the repressor to the various operator-binding sites.

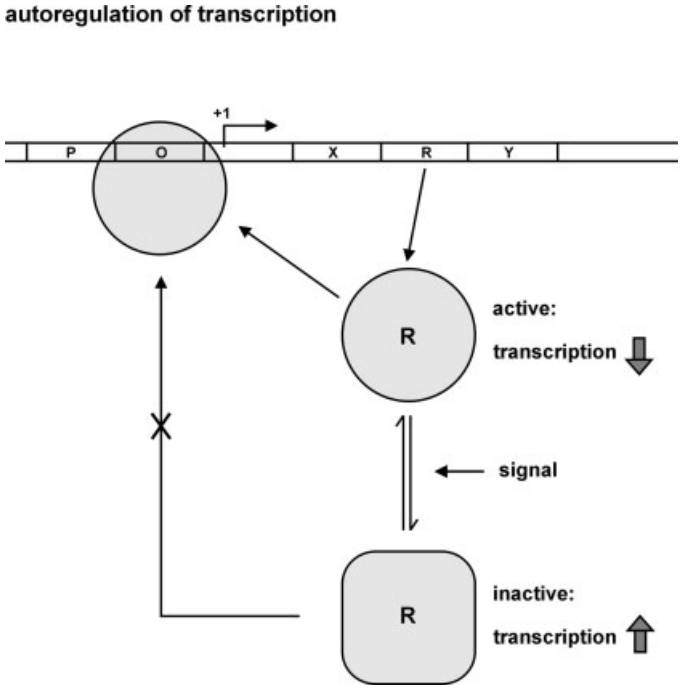


Fig. 1.16 Principles of autoregulation of transcription. In autoregulation, a repressor controls the transcription of its own gene as well as the transcription of other genes (X,Y).

An example for autoregulation is found in the hut-operon of *E. coli* and in the regulation of the SOS response in bacteria via the lexA repressor. There are examples of autoregulation at the level of translation as well (see Section 1.5.5.1).

With the aid of autoregulatory processes, it is possible for the cell to maintain a minimal concentration of repressor.

1.4 Regulation of Transcription in Eucaryotes

Procaryotes and eucaryotes differ decisively in the structure of the transcription start site and the complexity of the transcription apparatus. For a better understanding, we shall briefly summarize procaryotic transcription and then contrast it with eucaryotic transcription.

1.4.1 Overview of Transcription Initiation in Procaryotes

Transcription initiation in procaryotes is controlled via promotors and regulatory DNA sequences located near the promotor. The role of the promotor is to provide a defined association site for the RNA polymerase and to correctly orient it. The binding of the RNA polymerase to its promotor is controlled by the sigma factor, a component of the RNA polymerase holoenzyme. The sigma factor selects which genes are to be transcribed by specifically recognizing the promotor sequence and structure and by allowing the RNA polymerase to form a *transcription-competent complex* at the transcription start site.

Mechanism of Promotor Recognition

A transcription-competent complex must be present at the initiation site, with partial melting of the DNA, for the RNA polymerase to be able to add ribonucleotides complementary to the DNA template.

In a first approximation, the formation of a transcription-competent complex can be described according to a two-step mechanism (Fig. 1.17). The initial binding of the RNA polymerase to the promotor leads to the formation of a *closed complex* in which the RNA polymerase is only weakly bound. Isomerization of the closed complex transforms it into a transcription-competent open state. In the *open complex*, the RNA polymerase is tightly bound, and the DNA is partially unwound at the transcription start site.

The RNA polymerase of *E. coli* possesses with its subunit construction ($\alpha_2\beta\beta'\sigma$) a simple structure in comparison to eucaryotic RNA polymerases. The sigma factor is only required for the recognition of the promotor and the subsequent formation of a tight complex. After the incorporation of the first 8-10 nucleotides into the transcript, the sigma factor dissociates from the holoenzyme, and the remaining core enzyme carries out the rest of the elongation.

There are several sigma factors in *E. coli* (σ^{70}, σ^{54}, σ^{32}, σ^{28}) which can associate with the core enzyme to form the holoenzyme. The various sigma factors differ significantly with respect to their cognate promotor sequences. The overwhelming majority of the promotors in *E. coli* are recognized and activated by σ^{70}-containing holoenzymes.

Key elements of σ^{70}-dependent promotors are the TATA box, with the consensus sequence TATAAT 10 bp upstream from the transcription initiation site (pos. −10), and the sequence TTGACA at the position −35 (Fig. 1.18). Both sequences are necessary for the recognition of the promotor by σ^{70}. Structural analysis of the *Thermus aquaticus* RNA polymerase holoenzyme bound to DNA shows that all sequence-specific contacts with the core promotor are mediated by the sigma subunit (Murakami et al. 2002). This archaeal RNA polymerase has a subunit structure ($\alpha_2\beta\beta'\omega\sigma$) similar to that of the eubacterial enzyme. The intervening sequences, as well as other upstream sequences, can also influence the efficiency of transcription initiation. It is not possible to define consensus sequences at these positions. An optimal σ^{70}-dependent promotor can be defined as a sequence with the −35 hexamer as well as the −10 hexamer 17 bp away. The latter lies 7 bp upstream from the transcription initiation site.

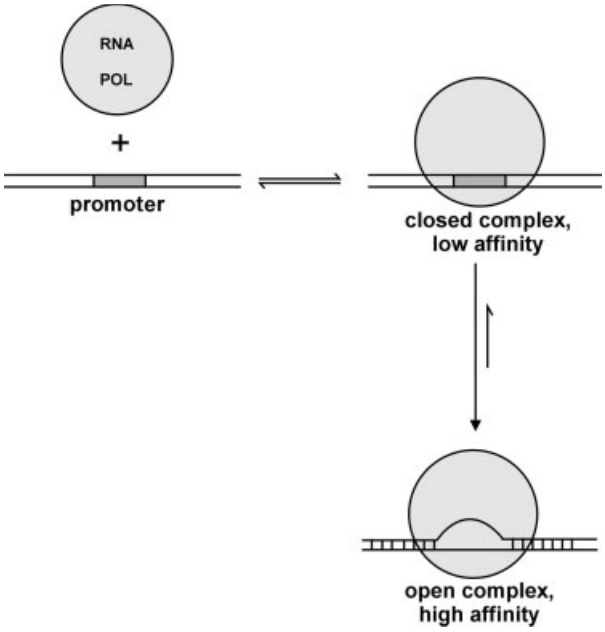


Fig. 1.17 Two-step mechanism of transcription initiation in procaryotes. The binding of a procaryotic RNA polymerase to its promotor can be subdivided into two steps. In the first step the RNA polymerase binds to the closed promotor with low affinity. The closed complex isomerizes in a second step to an open complex in which the promotor is partially unwound. Detailed consideration reveals that further steps can be distinguished. These are not shown here for simplicity reasons.

An important aspect of σ^{70}-dependent promotors is the fact that the holoenzyme is capable of initiating a weak transcription even without accessory proteins. In this way, the σ^{70}-containing holoenzyme can independently carry out all necessary steps, e.g., melting of the DNA, so that a constitutive transcription without the participation of regulatory proteins is possible. In this case, the extent of transcription depends on the affinity of the holoenzyme for the promotor and thus depends indirectly on the promotor sequence.

Regulatory DNA-binding proteins in procaryotic transcription

The recognition sequences of regulatory proteins may overlap not only the promotor site, but can also be found in the immediate vicinity of the σ^{70} promotor. The sequence elements are relatively simple and often include only one binding site for regulatory proteins.

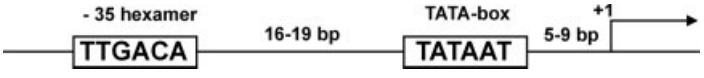


Fig. 1.18 Promotor structure of σ^{70}-dependent genes in *E. coli*.

Transcriptional activity is controlled mainly according to two mechanisms:

– *Repressors* turn off transcription by, for example, competing with the holoenzyme for binding to the promotor. RNA polymerases and repressors compete if the repressor-binding site overlaps with the promotor sequence.
– *Transcriptional* activators turn a gene on by increasing the efficiency of transcription initiation above the basal level. Transcription activation plays a particularly important role for the promotors classified as weak based on their sequence. It is assumed that the transcription activation occurs via protein-protein interactions between the DNA-bound transcriptional activator and the holoenzyme. This form of activation demands a close and defined contact. A change in the distance between the transactivator-binding site and the promotor may lead to a loss in the stimulatory effect if direct communication between the two proteins is no longer possible. On the other hand, a shift in the binding site of the transcriptional activators upstream in the direction of the promotor can lead to inhibition of transcription. Characteristic for the regulation of σ^{70}-dependent promotors by DNA-binding proteins is the tightly defined region within which the regulatory protein must bind relative to the binding site of the RNA polymerase holoenzyme.

Structural data on the interaction of transcriptional activators with the RNA polymerase holoenzyme are not yet available. We therefore rely on models to explain the mechanism of transcriptional activation. A plausible and experimentally supported model of transcriptional activation for σ^{70} promotors assumes that the transcriptional activators recruit the RNA polymerase holoenzyme to the promotor (Ptashne and Gann, 1997). According to this model, the DNA-bound transcriptional activator interacts with the holoenyzme, itself either free or DNA-bound, thus enabling the RNA polymerase to form a tight complex with the promotor. The recruitment increases the life span of the holoenyzme-DNA complex and/or eases the transition to the open, high affinity complex, from which transcription initiation occurs. Without transcriptional activators present, the holoenzyme binds only weakly to the promotor, and the transition to the open complex occurs only at a low frequency. Components of the holoenzyme that may be involved include the α-subunit or the σ-factor. Multiple, synergistically acting contacts can be formed between the transcriptional activator and the holoenzyme. This concept does not predict a conformational change in the holoenzyme, but does assume that the interactions between the holoenzyme and transcriptional activator are rather nonspecific in nature. As already proven experimentally, heterologous binding surfaces can also carry out this function.

1.4.2 The Basic Features of Eukaryotic Transcription

Eukaryotic transcription is a highly complex process where more than 100 different proteins collaborate in synthesizing RNA in a highly regulated manner. The basic components of the eukaryotic transcription apparatus have been characterized biochemically quite well, and its functions have been inferred from reconstitution experi-

ments using purified, often recombinant proteins. From the *in vitro* experiments, specific functions could be ascribed to distinct proteins, and for some of them structural information is available supporting the presumed biochemical function.

Based on these data, eucaryotic transcription has been shown to be dependent on the following components (Fig.1.19):

- ***RNA polymerases I, II and III*** carry the enzymatic activity for the synthesis of RNA on the DNA template and are composed of 10-12 subunits.
- ***General transcription factors*** (GTFs) help to localize the RNA polymerase correctly on the promotor and to form a transcription-competent initiation complex. They serve to impose a specific structure on the transcription start site, and some of them are required for elongation of the transcript.
- ***Specific transcription factors*** are sequence-specific DNA-binding proteins that mediate regulated transcription. They select the genes to be transcribed by binding to specific promotor or enhancer sequences, and they form activating or inhibiting contacts to the basal transcription machinery. The specific transcription factors receive signals for transcriptional regulation and transmit these signals to the basal transcription complex and to chromatin.
- ***Mediators*** are a class of proteins that mediate the contacts between specific transcription factors and the basal transcription complex. These proteins are found as part of RNA polymerase holoenzyme forms.

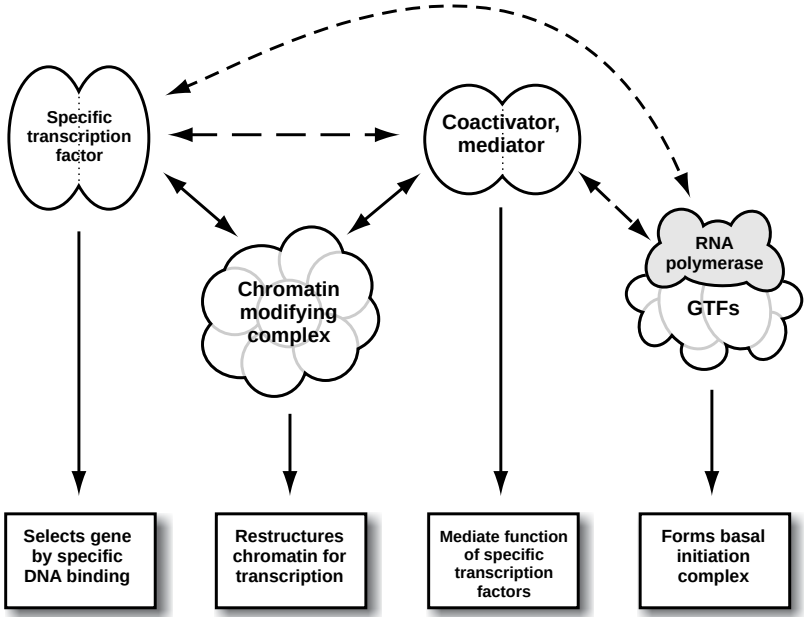


Fig. 1.19 Function and cooperation of the main components of eucaryotic transcription
GTF: general transcription factor.

- ***Chromatin-modifying and chromatin-remodeling activities*** are required for establishing a transcription-competent status on chromatin-covered promotors. The proteins involved are found in multi-subunit assemblies of varying composition. Specific transcription factors and mediators communicate with the chromatin-modifying activities.

1.4.3 The Eucaryotic Transcription Apparatus

Three types of RNA polymerases exist for the transcription of eucaryotic genes, each of which transcribes a certain class of genes. All three enzymes are characterized by a complex subunit structure.

RNA polymerase I is responsible for the transcription of the ribosomal RNA genes (class I genes), *RNA polymerase II* transcribes the genes encoding proteins (class II genes), and *RNA polymerase III* transcribes the genes for the tRNAs and the 5S ribosomal RNA (class III genes). Below, we shall limit the discussion to RNA polymerase II and the genes transcribed by it, since it plays the most important role for regulatory processes and signal transduction. Aside from this, many characteristics of the transcription of the genes of class II are also valid for genes of classes I and III.

1.4.3.1 Structure of the Transcription Start Site and Regulatory Sequences

Most promotors of of class II genes share three common features (Fig. 1.20A): the transcriptional start site, the TATA box, and sequences bound by transcriptional regulators. A typical core promotor containing the start site and the TATA box encompasses ca. 100 bp.

The TATA box or an initiation element are structural elements which define a minimal promotor required for recruiting the appropriate RNA polymerase and for initiation of transcription. In higher eucaryotes, the TATA box is often, though not always, ca. 30 bp from the transcription start site. The initiation element includes sequences in the immediate vicinity of the transcription start site. Not every eucaryotic promotor

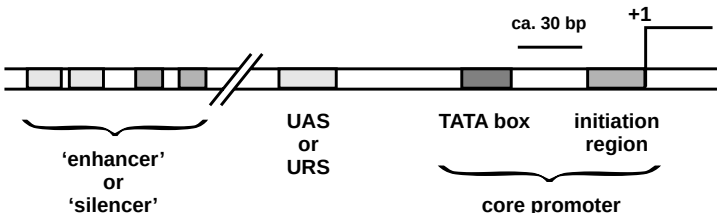


Fig. 1.20a Structure of a typical eukaryotic transcription start site. Enhancer elements and UAS or URS elements (UAS: upstream activating sequences; URS: upstream regulatory sequences) are binding sites for positive and negative regulatory DNA binding proteins. The TATA box is the binding site for the TATA box binding protein (TBP) and serves to position the RNA polymerase holoenzyme on the promoter. For promoters that do not possess a TATA box, this function is fulfilled by the initiation region.

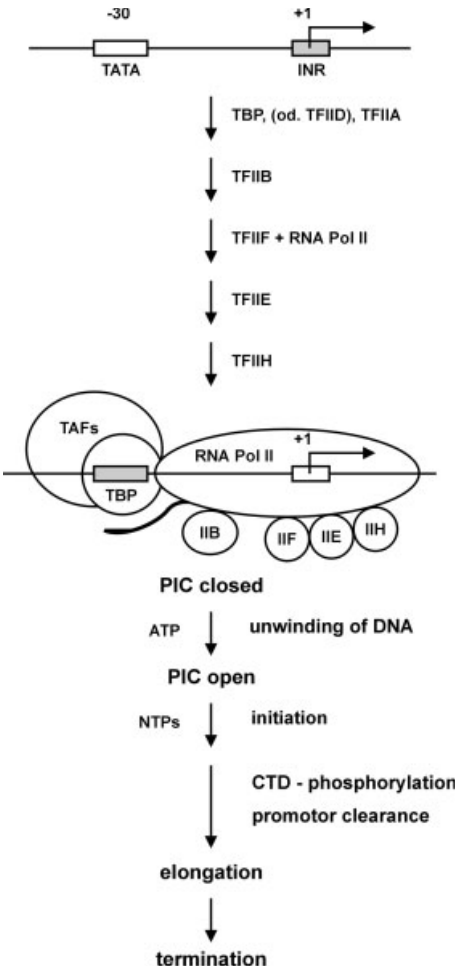


Fig. 1.20b Model for the formation of a pre-initiation complex and further steps for transcription initiation in eucaryotes. A functional pre-initiation complex (PIC) can be reconstructed *in vitro* from purified general transcription factors and RNA polymerase II. Successful reconstruction requires the addition of general transcription factors in the indicated order, whereby each of the given steps is supported by the experimentally confirmed existence of the implied intermediates. In the case of a TATA box-containing promoter, the first step in the reconstruction consists of the binding of the TATA box binding protein (TBP). It is assumed that under more physiological conditions the TFIID complex (together with associated TBP) binds to the promoter in the first step. TFIIA can join the pre-initiation complex after the binding of TBP at any of the following steps, but is not essential for the *in vitro* formation of a functional PIC. the pre-initiation complex is stabilized and favored by interactions of the general transcription factors with the TAFs (TAF: TATA box binding protein associated factor). After the formation of the closed pre-initiation complex and under ATP requirement, melting of the DNA and transition to the open complex occurs, in which the first nucleotide is added in the actual initiation step. The subsequent phosphorylation of the carboxyl terminal domain (CTD) of the large subunit of RNA polymerase II causes the dissociation of the RNA polymerase from the promoter and transition to the elongation phase. After termination the RNA polymerase is dephosphorylated and is ready for the next round of transcription (not shown in the figure). INR: initiation region.

possesses a TATA box. For promotors devoid of a TATA box, the initiation element is determining for promotor selection and formation of the pre-initiation complex.

The TATA box or initiation element are sufficient for the formation of a basal transcription apparatus composed of general initiation factors for transcription and RNA polymerase II (see Fig. 1.20B). Formation of the basal transcription initiation complex at the core promotor allows a basal transcription activity. A regulated transcription requires sequence-specific protein binding to regulatory sequences that can be be proximally or distally located. Regulatory sequences proximal to the core promotor are called Upstream Activating Sequences (UAS) or Upstream Repressing Sequences (URS), depending on whether the bound protein is an activator or a repressor of transcription.

Regulatory sequences can also be located far from the promotor and are called *enhancers* if the cognate DNA-binding protein is a transcriptional activator. They influ-

ence transcription independently of their orientation and at distances as great as 85 kb from the start site. Enhancers typically contain clusters of DNA-binding sites, and transcription activation results from the complex concerted action of various specific DNA-binding proteins. A multiprotein complex is formed at enhancers, called the enhanceosome (see Section 1.4.7.3).

Sequence elements that can repress transcription in an orientation- and position-independent fashion are called *silencers*.

1.4.3.2 **Elementary Steps of Eucaryotic Transcription**

As in procaryotes, the elementary steps of initiation, elongation and termination can be distinguished in eucaryotic transcription. Aside from the specific RNA polymerases, transcription in eucaryotes requires the action of numerous other proteins, which include transcription factors, mediators or coactivators and chromatin-modifying and -restructuring components. Transcription factors are required at the level of initiation, elongation and termination, and are accordingly known as initiation factors, elongation factors and termination factors of transcription.

Transcription in eucaryotes can, as shown schematically in Fig. 1.20B, be subdivided in the following steps:

- **Formation of a pre-initiation complex**
  Promotor selection, binding of basal initiation factors
  Binding of the RNA polymerase II
  Formation of a basal transcription apparatus
- **Activation of the pre-initiation complex**
  Melting of the DNA in the vicinty of the start site
- **Initiation**
  Incorporation of the first nucleotide
- **Transition from initiation to elongation**
  Processive RNA synthesis by RNA polymerase
- **Termination**
  End of synthesis at defined sequence elements.

Transcription is regulated to a great extent at the start of transcription, i.e. steps (1 – 4). These steps are thus the central point of the following discussion. One must keep in mind that *in vivo* transcription does not occur on naked DNA, but rather on chromatin, i.e. nucleosome-coated DNA. Activation of transcription requires an active remodeling of the chromatin structure at the transcription start sites and within the transcribed regions. Every discussion of the individual steps of transcription must consider this fact. However, only incomplete *in vitro* systems are available for the study of transcription of chromatin-associated DNA, so that in the following discussion only the data on the transcription of naked DNA is presented.

1.4.3.3 Formation of a Basal Transcription Apparatus from General Transcription Factors and RNA Polymerase

In contrast to the procaryotes, where the σ^{70}-holoenzyme of the RNA polymerase can initiate transcription without the aid of accessory factors, the eucaryotic RNA polymerase requires the help of numerous proteins to begin transcription. These proteins are termed *basal* or *general initiation factors of transcription*. Together with RNA polymerase II, they participate in the basal or core transcription apparatus. Reconstitution experiments starting from naked DNA and purified general transcription factors have shown that the various components associate in a defined order for the formation of a transcription-competent complex (Fig.1.20B), from which a low level of transcription is possible. An increase in the basal transcriptional level requires the effect of specific transcriptional activators which bind cognate DNA sequences at a variable distance from the promotor. With purified GTFs and RNA polymerase alone, however, the response of transcriptional regulators observed *in vivo* cannot be reconstituted fully. In the living cell, the DNA template is packaged into chromatin, which adds another level of complexity to transcriptional regulation. It has been clearly shown that activators or repressors can help to reorganize and modify chromatin to establish a transcriptionally active or inactive state. The transcriptional regulators thus interact both with the basal transcription machinery and with the protein apparatus that is involved in chromatin modification and reorganization.

A transcription-competent pre-initiation complex consisting of general transcription factors and RNA polymerase II can be reconstituted in the test tube from the individual components. As outlined in Fig. 1.20B, efficient reconstitution requires a defined order for the addition of the individual components.

Tab. 1.1 General initiation factors of transcription by RNA polymerase II. TAF: TATA box binding protein associated factor; TBP: TATA box binding protein; RNA Pol II: RNA polymerase II.

protein	number of subunits	subunit size (kDa)	function
TFIID:			
TBP	1	38	sequence specific binding to TATA box, recruitment of TFIIB
TAFs	12	15–250	promoter recognition, regulation, chromatin modification
TFIIA	3	12, 19, 35	stabilization of TBP-DNA binding; antirepression
TFIIB	1	35	recruitment of RNA Pol II–TFIIF; selection of start site by RNA Pol II
TFIIF	2	30, 74	assists in promoter binding by RNA Pol II
RNA polymerase II	12	10–220	enzymatic activity of RNA synthesis, binding of TFIFF
TFIIE	2	34, 57	binding of TFIIH, modulation of activities of TFIIH
TFIIH	9	35–89	helicase, protein kinase and ATPase-activity; promoter unwinding, promoter clearance (?)

The basal transcription factors (summarized in Table 1.1) can be characterized in the following terms:

- **TFIID, TAFs and TBP**
  The transcription factor TFIID is a multi-protein complex that binds specifically to the promotor region. It consists of the TATA box-binding protein (TBP) and TATA box-binding protein associated factors (TAFs). Recognition of the core promotor can be mediated by two components of TF IID, depending on promotor structure (see Fig. 1.21). In TATA box-containing promotors, TBP specifically recognizes the TATA box, and its binding leads to a distinct bending of the DNA (see Fig. 1.9). In this manner, a particular topology of the DNA is created that serves as a prerequisite for the defined binding of further basal transcription factors, such as TFIIA and TFIIB (review: Tan and Richmond, 1998). For promotors lacking a TATA box, the initiation element serves as a specific contact point for binding by the TAF components TAF_{250} and TAF_{150}.
  It is assumed that the binding of TFIID via TBP to the TATA box represents an important regulatory step in the recognition and selection of the promotor *in vivo*. The TAFs comprise at least 12 different proteins that fulfill numerous functions (review: Struhl and Moqtaderi, 1998, Albright and Tjian, 2000). On the one hand they are ascribed a structure-promoting function. Some of the TAFs display a high degree of homology to the histones H2A, H3 and H4. Their structure matches the canonical histone-fold dimer, and TAF-dimers are formed via the histone-fold. It is therefore speculated that TAFs impose a distinct topology to the DNA and help to create a nucleosome-like structure at the promotor.
  Furthermore, the TAFs are targets for protein-protein interactions with transcriptional activators and two of them are required for sequence-specific binding to the initiaion element of TATA-less promotors. TAFs also possess enzymatic activity. $TAF_{II}250$ has both a histone acetylase activity and a protein kinase activity. While the former presumably plays a role in the reorganization of the nucleosome, the latter can lead to phosphorylation of TFIIF.

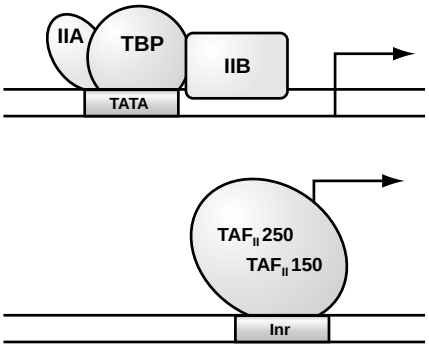


Fig. 1.21 Initial protein complexes formed at the TATA box and the initiation region
Selection of the transcription start site is mediated either by the TATA box or the initiation region INR. The TATA box is bound by the TATA box binding protein TBP in cooperation with the general transcription factors IIA and IIB. In the absence of a TATA box, transcription initiation requires binding of the $TAF_{II}250$ and $TAF_{II}150$ to the INR.

It has also been shown that the composition of TFIID is not fixed, but may vary depending on the detailed structure of the promotor (review: Veenstra and Wolffe, 2001).

- **TFIIA and TFIIB**
  TFIIA and TFIIB support TFIID in the formation of a stable complex with the promotor. TFIIB is necessary for the downstream selection of the start site for RNA polymerase II. Interactions with TFIIB ensure correct positioning of the RNA polymerase II on the promotor. Crystal structures have been solved for several of the intermediates of the pre-initiation complex (review: Werner and Burley, 1997), showing, for example, that TBP effects a predominant kink in the DNA (see Fig.1.9). TFIIB binds to the TBP-DNA complex, contacting both TBP and the DNA.
- **TFIIF, TFIIE**
  TFIIF is found in a pre-formed complex with RNA polymerase II and suppresses the nonspecific binding of RNA polymerase to DNA. TFIIF supports the association of RNA polymerase with the promotor-bound complex of TFIIA, TFIIB and TFIID. TFIIE binds TFIIH to assist the latter with the melting of the promotor.
- **TFIIH**
  The binding of TFIIH completes the formation of the pre-initiation complex. TFIIH is a multi-protein complex with a variable composition (see Section 1.4.3.5) and which possesses protein kinase, ATPase and helicase activities. The helicase activity of TFIIH is required for the melting of the promotor.

Overall, the general transcription factors can be assigned the role fulfilled by a single protein in procaryotes – namely the σ-factor. This role includes the correct positioning of the RNA polymerase on the promotor and the preparation for the incorporation of the first nucleotide. The addition of ATP to the pre-initiation complex leads to a rapid melting of the promotor, initiation of RNA synthesis, and dissociation of the RNA polymerase from the promotor.

Holoenzyme Forms of RNA Polymerase II

RNA polymerase II can be isolated from the cell in various forms. From yeast and from metazooans, holoenzyme forms of RNA polymerase have been purified that are composed of the core of RNA polymerase II and other accessory transcription factors, suggesting that a preassembled transcription apparatus can be assembled at the promotor in one step. These holoenzymes are formed from the core of RNA polymerase associated to a variable degree with the general transcription factors and with one or more other subunit complexes called mediators or coactivators.

The core of the RNA polymerase II is a functional unit composed of 12 subunits. The structure of the core of RNA Pol II from yeast has been resolved, providing a comprehensive model of template- and mRNA-binding as well as of the location of the active site of nucleotide incorporation (Cramer et al., 2001).

Holoenzyme forms of RNA polymerase II have general transcription factors associated. TFIIB, TFIIE, TFIIF and TFIIH have been identified as components of the RNA polymerase II holoenzyme of yeast.

The yeast holoenzyme contains further proteins, known as *mediators* or *SRB proteins* (SRB, suppressor of RNA polymerase B). The mediators function as coactivators (see Section 1.4.4.2) and appear to integrate signals from transcriptional activators at promotors. A complete structural characterization of the holoenzyme has proven to be difficult, because some of the proteins accessory to the core enzyme are not permanently and are often only loosely associated with the core of RNA polymerase II. Clearly different forms of the RNA polymerase II holoenzyme exist in the living cell, each of slightly different composition and function.

The use of a pre-formed holoenzyme complex of the RNA polymerase II appears to be an economical mechanism for the formation of an initiation-competent transcription complex. There are two basic processes involved in the formation of a pre-initiation complex in the cell:

– promotor recognition and promotor binding by TFIID (TBP and TAFs)
– binding of the RNA polymerase holoenzyme to the promotor-bound TFIID.

Transcriptional activators or repressors can intervene as regulators at various steps in the initiation of transcription. They can interact with components of TFIID, as well as with components of the RNA polymerase II holoenzyme to stimulate transcription. Furthermore, contacts to chromatin proteins are formed to change the structure of chromatin during transcription initiation and ongoing elongation. Regulated transcription and chromatin modification generally requires the aid of further protein components, which are commonly termed *coactivators* (see Section 1.4.4.2).

1.4.3.4 **Phosphorylation of RNA Polymerase II and the Onset of Transcription**

Phosphorylation of the large subunit of RNA polymerase II plays an important role at the beginning of the transcription process and during elongation (review: Kobor and Greenblatt, 2002). The large subunit of the mammalian enzyme contains 52 copies of the heptamer sequence YSPTSPS in the C-terminal domain (CTD) at which phosphorylation occurs. Phosphorylation occurs extensively on the Ser-residues of the CTD, to a lesser degree at the Thr-residues, and, very rarely, at the Tyr-residues. Two forms of RNA polymerase II can be isolated from cellular extracts: an underphosphorylated form and a hyperphosphorylated form. The isoforms fulfill different functions: RNA polymerase found in the initiation complex tends to display little or no phosphorylation at the C-terminus of the large subunit, while RNA polymerase II active in elongation is hyperphosphorylated in this region of the protein.

Based on the above observation, phosphorylation at the C-terminus is believed to serve as a trigger for the transition from initiation to elongation (Fig. 1.22) by exchanging cofactors bound to the CTD. The unphosphorylated C-terminus forms contacts with the TATA box-binding protein as well as with the SRB/mediator complex (see below). The high density of negative charges at the C-terminus resulting from phosphorylation is assumed to disrupt these interactions, thereby releasing RNA polymerase into the elongation process and allowing the association of other cofactors.

The association of transcriptional cofactors with the CTD of elongating RNA polymerase II apparently plays an important role in the maturation of the pre-mRNA and

in the formation of the mRNA processing factory. For this reason, the CTD is associated with components of the splicing apparatus, including proteins with high homology to splicing regulatory proteins like SR-Proteins (see Section 1.5.2). Furthermore, proteins involved in the capping of the mRNA and in polyadenylation associate with the CTD during transcription elongation, establishing a firm link between transcription and mRNA processing (review: Proudfoot et al., 2002).

Three different protein kinases have been identified that can phosphorylate the CTD. One protein kinase involved in phosphorylation at the C-terminus is localized in TFIIH. It has been shown that TFIIH contains a Ser/Thr-specific protein kinase termed CDK7 (or MO15). CDK7 belongs to the family of cyclin-dependent protein kinases (see Section 13.2.1). The cognate cyclin, cyclin H, is also found in TFIIH. CDK7 is identical to the CDK-activating protein kinase CAK, to which is ascribed an important role in the regulation of the cell cycle. The functional importance of this identity is not yet fully understood.

The Srb/mediator complex contains another protein kinase/cyclin pair, namely CDK8 and cyclin C, which participates in phosphorylating the CTD. The third protein kinase capable of CTD phosphorylation is CDK9/Cyclin T, which is the target of transcription activation by the retroviral TAT protein (review: Romano et al., 1999). Altogether, CTD phosphorylation has proven to be a point where many regulatory signals may converge and influence the transition from initiation to elongation, the efficiency of elongation and the maturation of the mRNA. Mechanistic details on how the various kinases cooperate and how they are regulated are however still lacking.

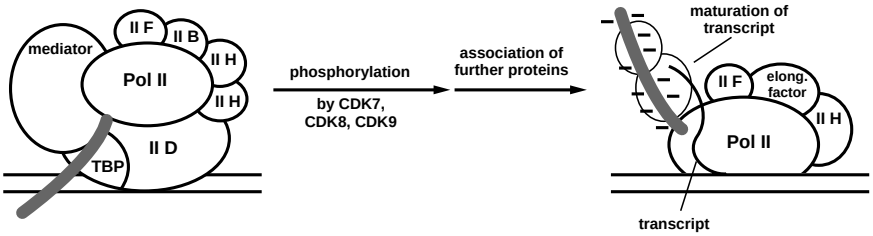


Fig. 1.22 Phosphorylation of the C-terminal domain of RNA polymerase II and the beginning of transcription. The transition from the initiation complex to actual begin of transcription is regulated via phosphorylation of the C-terminal domain (CTD) of RNA polymerase II. In the above model it is assumed that initially a complex is formed between TFIID and a holoenzyme of RNA polymerase consisting of RNA polymerase II and associated factors (mediators, SRB proteins) and the basal transcription factors. Phosphorylation of the C-terminal domain - and of other subunits of the holoenzyme - effects the dissociation of the RNA polymerase from the initation complex, the association of further protein factors and the transition to the elongation phase. The cyclin-dependent protein kinases CDK7, CDK8 and CDK 9 have been implicated in these phosphorylations. The nature of the signal that induces phosphorylation of RNA polymerase II remains unknown. SRB: suppressor of RNA polymerase B.

1.4.3.5 TFIIH – a Pivotal Regulatory Protein Complex

TFIIH is a multi-protein complex consisting of eight different subunits (review: Egly, 2001), which can be separated into two subcomplexes: the core TFIIH complex and a kinase/cyclin subcomplex (see Fig. 1.23).

Three enzymatic activities are found in TFIIH:

– DNA-dependent ATPAse
– ATP-dependent helicase
– CTD-phosphorylating protein kinase.

TFIIH is required for critical, early steps of transcription, including the transition from initiation to elongation. In addition to its function in phosphorylating the CTD, TFIIH opens the DNA template by its ATP-dependent helicase activities located on the XPB and XPD proteins. Electron microscope studies have revealed a ring-like structure of TFIIH, with a hole whose size is able to accommodate a double-stranded DNA (Chang and Kornberg, 2000).

TFIIH also participates in another important cellular function, namely nucleotide excision repair of damaged DNA. This function accounts for the observation that transcription and the removal of bulky base adducts by nucleotide excision repair (NER) are coupled. An increased repair of DNA damage by NER is observed while a gene is being transcribed. During transcription-coupled repair, TFIIH assembles with other repair proteins into a large repair complex, allowing for the removal of DNA adducts.

The presence of the protein kinase CDK7 and of Cyclin H in the kinase subcomplex indicates a link between transcription and cell cycle regulation. CDK7 is identical to cyclin-activating kinase, CAK (see Section13.2.1), which regulates cell cycle transitions.

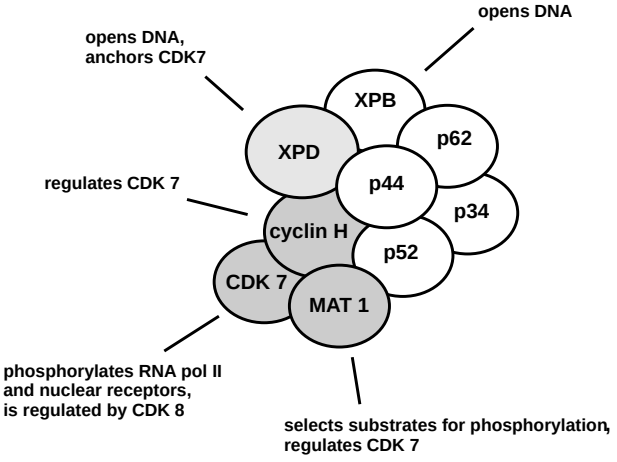


Fig. 1.23 The subunits of TFIIH and their presumed functions. CDK: cyclin dependent protein kinase; XPB, XPD, xeroderma pigmentosum subgroups B, resp. D complementing proteins with helicase activity.

TFIIH itself, or other individual components of TFIIH, thus participate in the following fundamental processes in the cell:
– transcription
– nucleotide excision repair of DNA lesions
– regulation of the cell cycle.

The picture of the structure and function of TFIIH is thus varied and complex. TFIIH or components of it can assemble into different multiprotein complexes which perfom central functions in the cell. The mechanistic details and the regulation of the various activities have long been not understood.

1.4.4 Regulation of Eucaryotic Transcription by DNA-binding Proteins

Primary controlling elements of the transcriptional activity in eucaryotes are specific DNA-binding proteins. They bind cis-acting DNA elements and have a specific influence on the initiation of transcription.

As in procaryotes, there are transcription-activating proteins (the transcriptional activators), as well as proteins which inhibit transcription (the transcriptional repressors). Both types of regulatory proteins can influence transcription in two ways:
– via the RNA polymerase holoenzyme
– via the chromatin structure.

The activators of transcription have been studied and characterized the most extensively and will be discussed first.

1.4.4.1 The Structure of Eucaryotic Transcriptional Activators

It is generally true that eucaryotic transcriptional activators act via direct or indirect protein-protein contacts with the transcription apparatus and/or chromatin components. In this context, the definition of the transcription apparatus is rather broad: it includes the basal apparatus as well as proteins which act as coactivators or mediators.

For a DNA-binding protein to engage in the regulation of activation of transcription it must possess the following functions:
– specific DNA binding
– communication with the transcription apparatus and/or chromatin
– the ability to be regulated by effectors.

In many cases the various functions are located on independently folding protein domains, resulting in modularly constructed transcriptional activators.

DNA Binding

Transcriptional activators bind specifically to cognate DNA elements variably located relative to the promotor and can interact directly or indirectly with the transcription apparatus. Transcriptional activators depend on the occurrence of regulatory DNA elements for their action and perform their function on specific genes. They are thus also termed specific transcriptional activators, in contrast to proteins called coactivators that activate transcription independently of specific DNA elements (see Section 1.4.4.2).

The DNA binding serves to bring the binding protein into proper orientation to the pre-initiation complex. Binding to *cis*-elements creates a high effective concentration of binding proteins close to the transcription apparatus to enable productive interactions. Because protein-induced conformation changes of the DNA aid in orienting the transcription complex on the promotor, the conformation and bending of DNA also plays an important role. The structural motifs of eucaryotic transcriptional activators for the recognition of specific DNA sequences are varied and include all the DNA-binding motifs summarized in Section 1.2.1.

In many cases, the specific DNA binding serves only to create a high concentration of the transcription factor in the vicinity of the site of transcription initiation. This property can be used in domain-swapping experiments to alter the specificity of regulation by design. Experiments such as the one described in Fig. 1.24 have contributed significantly to our understanding of the structure and function of eucaryotic transcriptional activators. A prerequisite for successful *domain-swapping experiments* is that the DNA-binding domain should not influence the function of the transactivating domain. This is often, but not always, the case.

Communication with the Transcription Apparatus

The first data concerning the structural requirements for communication with the transcription apparatus came from domain-swapping experiments with the GAL4 protein of yeast.

One structural domain in the GAL4 protein could be shown to mediate a stimulation of transcription and was thus termed the *trans-activating domain*. The essential structural element of the *trans*-activating domain of GAL4 is an amphipathic α-helix with negatively charged and nonpolar amino acid side chains on opposite sides of the helix. The amount of negative charge correlates with the extent of the *trans*-activating effect and also determines the distance over which an effect on the transcription complex is still possible.

Examples of other *trans*-activating domains are the glutamine-rich domains of the transcription factor Sp1 and the proline-rich domain of the transcription factor CTF/NF1, which contains 20% proline residues.

It is assumed that *trans*-activating domains are structural elements that can adapt to become complementary to a surface of the transcription apparatus in a flexible and rather unspecific manner. Furthermore, parts of the *trans*-activating domain mediate contacts to the chromatin-remodeling and chromatin-modifying complexes and recruit these to the enhancer or promotor region. Information on the structure of

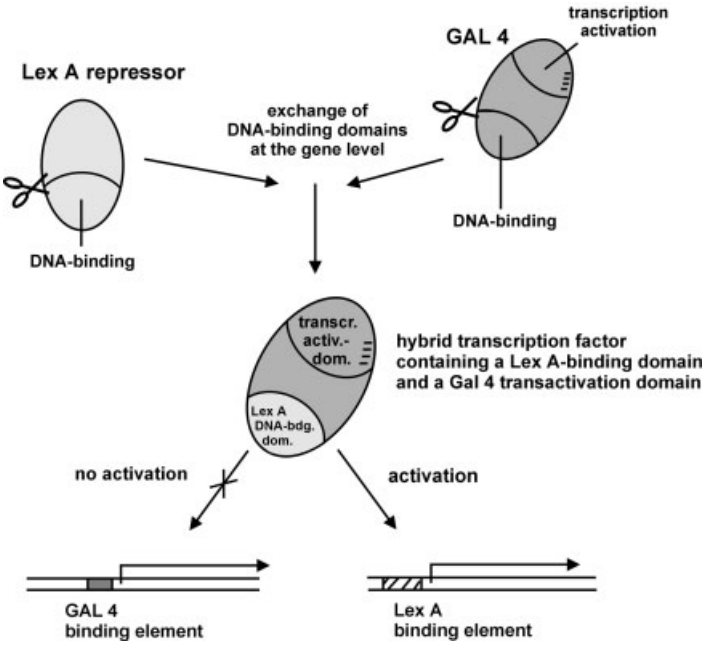


Fig. 1.24 "Domain Swapping" experiment for the creation of hybrid transcription factors. The functional domain of transcription factors often behave as independent structural units that can be be exchanged among various proteins by genetic means. In the above example of a domain swap experiment, the DNA binding domain of the bacterial LexA repressor is transferred to the GAL4 protein of yeast. GAL4 is a transcription activator that positively regulates the transcription of the enzymes of galactose metabolism. The genetically engineered hybrid transcription factor has the DNA binding specificity of the LexA repressor and the transactivating specificity of the GAL4 protein. The hybrid transcription factor activates promotors with the LexA binding element, but not promotors with the GAL binding element.

the *trans*-activating domain, as well as on the overall protein structure, is currently available for the steroid hormone receptors, among others (see Chapter 4).

1.4.4.2 **Concerted Action of Transcriptional Activators and Coactivators in the Regulation of Transcription**

The formation of an active, regulation-competent initiation complex for transcription in eucaryotes demands the concerted action of a large number of proteins. It is estimated that more than 100 different proteins participate in the initiation of transcription in eucaryotes. The basal transcription complex, consisting of the general initiation factors, as well as RNA polymerase II, only allows for a slow transcription rate. For a regulated acceleration of this low transcription rate by transcriptional activators, further proteins named coactivators or corepressors are required. The terms coactivator or corepressor are miscellaneous terms referring to proteins that are required for the functions of transcriptional activators (or repressors) but are not required for

basal transcription. They do not show site-specific DNA binding by themselves and mediate the regulatory function of the site-specific DNA-binding proteins bound at the promotor or at enhancers. Many functions of the coactivators or corepressors are linked to the processes of chromatin restructuring and chromatin modification.

The regulated activation of transcription thus requires at least two types of proteins (Fig. 1.25):

- **Specific transcriptional activators**
  The specific transcriptional activators (see Section 1.4.3.1), represented by the GAL4 protein of yeast, are sequence-specific DNA-binding proteins that bind to regulatory DNA sequences. They possess both a DNA-binding domain and a *trans*- activating domain to allow them to interact directly with the transcription apparatus.
- **Coactivators**
  Extensive studies on the reconstitution of a regulated transcription initiation in *in vitro* systems have shown that most of the specific transcription factors are not capable of stimulating transcription above the basal level without the assistance of further proteins. Further coactivators are required for this task. The coactivators may be subdivided into three classes (Fig. 1.26).

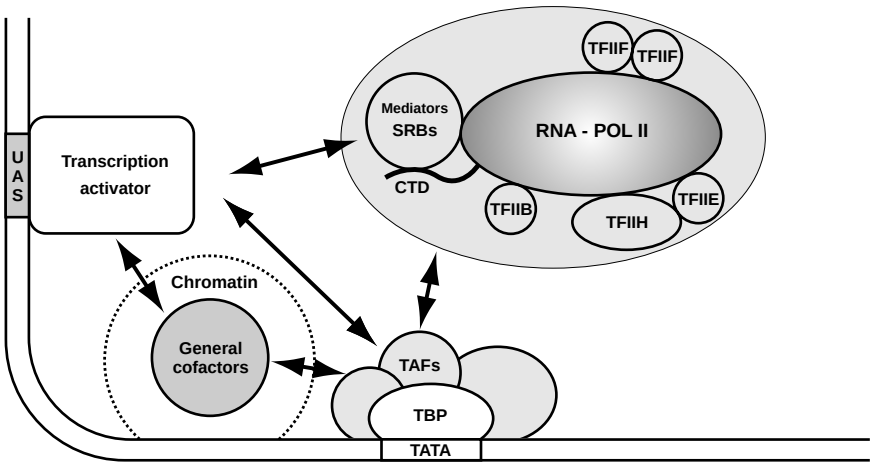


Fig. 1.25 Activators and coactivators of transcription intiation. The figure shows the function of three groups of proteins that function as coactivators. The general cofactors mediate the interactions between the specific transcription activators and the TFIID complex as well as with various forms of the RNA polymerase II holoenzyme. The TAFs are components of the TFIID complex and serve as contact points for specific transcription activators. The mediators are found as large protein complexes associated with various forms of holoenzymes of RNA polymerase II. SRB proteins belong to the class of mediators, which, among other things, interacts with the CTD of RNA polymerase. The simplified diagram does not show the interactions with chromatin.
TBP: TATA box binding protein, TAF: TATA box binding protein associated factor, SRB: suppressor of RNA polymerase B, CTD: C-terminal domain of RNA polymerase II, UAS: upstream activating sequence.

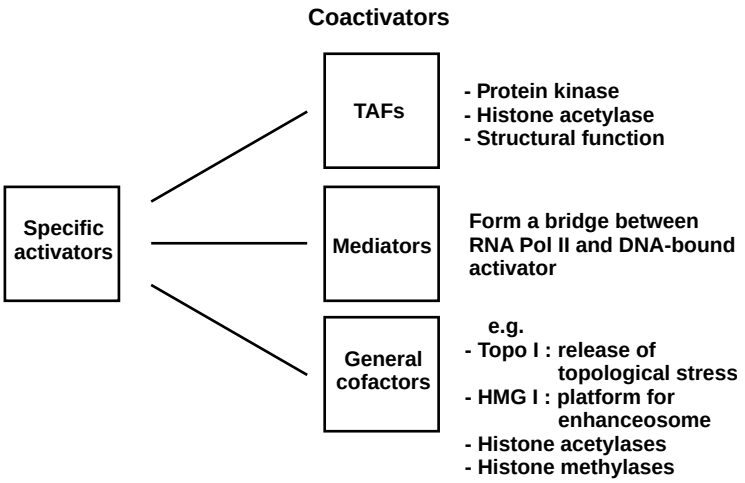


Fig. 1.26 The main classes of transcriptional coactivators. Specific transcription factors form functional interactions with coactivators classified as TAFs, mediators and general cofactors.
Topo I: DNA topoisomerase I; HMG: high mobility group protein.

TATA Box-binding Protein Associated Factors (TAFs)

The TAFs are components of TFIID (see Table 1.1) and are required for a regulated transcription (review: Albright and Tijan, 2000). Thus, the stimulation of transcription by the transcriptional activators Sp1 and NTF-1 depends upon the presence of specific TAFs in the TFIID complex. The TAFs mediate interactions between the transcriptional activators and the TFIID complex; in many cases direct protein-protein interactions could be demonstrated between the activators and TAFs. Some of the TAFs possess enzymatic activities which allow them to participate in the regulation of transcription. By this token, the histone acetylase and protein kinase activity of $TAF_{II}250$ is ascribed a regulatory function in the remodeling of chromatin and in the control of the activity of the basal transcription factors.

Overall, the physiological function of individual TAFs is still incompletely understood. The situation is complicated by the fact that a subset of the TAFs, including the histone-like TAFs, have been identified as components of a large histone acetylase complex, termed SAGA in yeast and PCAF (p300/CBP associated factor) in humans (see Section 1.4.7).

Mediators/SRBs

Mediators include proteins which are organized in multiprotein complexes as a conserved interface between the gene-specific transcription activators and the basal transcription apparatus. They transduce positive or negative regulatory information from promotors or enhancers to RNA polymerase II, modulating its activity. Mediators include, among others, the SRB proteins (see Section 1.4.3.3). A distinct mediator complex from yeast has been described which comprises 20 proteins and is found either in

a free state or as part of the RNA polymerase II holoenzyme (review: Myers and Kornberg, 2000). Some of the mediator subunits interact with transcriptional activators and some contact the CTD of RNA polymerase II. Examples of metazoan mediator complexes are the TRAP complex, the PC2 complex and the SMCC complex (review: Malik and Roeder, 2000).

General Coactivators

Besides the TAFs and the mediator/SRB complex, other coactivators have been described which are required for the function of transcriptional activators either in a general sense or for specific activators only.

Among the rather loosely defined class of general coactivators are several proteins which bind DNA and associate with the chromatin. Some possess enzymatic activity. Examples are DNA-topoisomerase I, poly-ADP-ribose polymerase, and HMG1 and HMG2 proteins (HMG: "high mobility group"), both associated with the chromatin (review: Thomas and Travers, 2001). Of particular interest is HMG1(Y), which has now been identified as an essential structural component of the enhanceosome (see Section 1.4.7.3).

Among this type of coactivators are also included proteins with histone acetylase activity (HAT). The acetylation/deacetylation of histones plays a significant role in transcription regulation of chromatin-coated DNA. Coactivators with histone acetylase activity are the CBP protein, the p300 protein, and the GCN5 protein (see Section 1.4.7).

Experiments on the reconstitution of regulated transcription suggest a cooperative formation of a regulation-competent transcription apparatus. All three classes of coactivators appear to be necessary for an efficient and regulated transcription event. Most of the coactivators are organized in multiprotein complexes that can operate in a modular fashion. Depending on the specific demands of the gene to be regulated the composition of the coactivator complexes may vary as well as their association with the RNA polymerase core enzyme. The presence of one particular transcriptional activator and its *cis* DNA element clearly is not sufficient for transcription activation. Generally, the TAF proteins, the mediators, as well as the general cofactors must be present in order for the regulatory signal to be transmitted to the basal transcription apparatus. In addition, the coactivators are ascribed a structural role by imposing a specific structure at the transcription start site that is required for efficient transcription initiation. The structural functions of coactivators are specifically required during transcription of chromatin-covered DNA. In this situation, which most closely resembles transcription in the living cell, another layer of regulation is required in order to overcome repressive effects of DNA packaged into chromatin. Therefore, restructuring and modifying chromatin is another major aspect of regulated transcription (see Section 1.4.7). The availability of the coactivators can, as with the transcriptional activators, be controlled in a tissue-specific manner either via gene expression or via signal transduction chains. This opens up a realm of possibilities for tissue-specific transcription activation.

1.4.4.3 Interactions with the Transcription Apparatus

The activating domain of transcriptional activators requires specific binding partners within the transcription apparatus. Unequivocal identification of these binding partners has proven difficult because of the large number of proteins that participate in the formation of the pre-initiation complex.

Furthermore, the target surfaces of the binding partners appear to be not as specific as one would intuitively assume. Much structural evidence suggests that activation domains expose flexible hydrophobic elements to contact hydrophobic patches on the target. Since these interactions are *per se* not highly specific, the colocalization of the activation domain and the target protein on the DNA template seems to be an important factor for activation. In addition, cooperative interactions of transcription factors bound to multiple sites at enhancers may help to increase the efficiency of contact formation with the pre-initiation complex. In these interactions, multiple targets may be involved. The targets can belong, e.g., to the group of basal transcription factors, such as TFIIB, and/or to the group of coactivators, such as the TAF proteins of the TFIID complex, proteins of the mediator complex and proteins required for chromatin restructuring.

The pathways by which specific transcriptional activators and coactivators influence transcription initiation are still a matter of speculation. There are two main pathways under discussion, which possibly act cooperatively:

In one model it is assumed that transcriptional activators and coactivators increase the efficiency of formation of the pre-initiation complex. This function includes a relief of the repressive structure of the chromatin by restructuring of the chromatin at the transcription start site.

The other model views the activators and coactivators as responsible for the stable and defined spatial arrangement of proteins in the holo-complex and for the induction of topology in the DNA which promotes initiation.

It can be assumed that the extent to which either pathway is used depends on the structure of the specific gene, as well as on the structure of the chromatin.

1.4.5 Regulation of the Activity of Transcriptional Activators

The repertoire of mechanisms for control of the activity of eucaryotic transcriptional activators (and also of coactivators) is varied and allows a spatially and temporally coordinated regulation of transcription.

The principle means by which the activity of sequence-specific DNA-binding proteins is controlled have already been presented in Section 1.3. The importance of these mechanisms for regulation in eucaryotes will be discussed below. Altogether, the demands on eucaryotic organisms with regard to the regulation of transcription activity are much more complex than those on procaryotes. This tenet holds for the structure of the transcription apparatus as well as for the mechanism of transcription regulation.

1.4.5.1 The Principal Pathways for the Regulation of Transcriptional Activators

Figure 1.27 gives an overview of the most important mechanisms by which the transcription-regulating activity of specific DNA-binding processes in eucaryotes can be controlled. They include de *novo synthesis*, as well as the *modification* and *availability* of *pre-existing proteins*. The specific expression of transcriptional activators is of great importance during the development and differentiation of organisms where long-term changes in gene expression are required. *Concentration gradients* of diffusible regulatory proteins for the specific control of gene expression are used in development, as shown for the bicoid protein of *Drosophila* (Driever and Nuesslein-Vollhard, 1989).

For the coordination of metabolic processes or for the regulation of cell division, the cell relies primarily on post-translational modification of pre-existing regulatory proteins. The activity or availability of pre-existing regulatory proteins can be adjusted rapidly and effectively by post-translational modifications, so that an immediate reaction within the framework of intercellular communication is possible. The modification of pre-existing regulatory proteins thus plays a more important role amongst eucaryotes than in bacteria. A further significant regulation mechanism is the *binding of effector molecules* to the DNA-binding proteins (compare Section 1.3.3). The influence of steroid hormones on the DNA-binding activity of steroid hormone receptors is an example of this regulatory mechanism.

Other important issues in the control if transcriptional regulators are the subcellular localization as well as the targeted degradation. We know of numerous examples of transcription factors that are controlled by a signal-dependent translocation from the cytoplasm to the nucleus or vice versa (see below, Section 1.4.5.2). Furthermore, specific signals can induce the degradation of a transcription factor via the ubiquitin-proteasome pathway (see Section 2.6) and thus specifically weaken the transcriptionally regulatory signal.

1.4.5.2 Phosphorylation of Transcriptional Activators

The phosphorylation of proteins on Ser, Thr or Tyr residues is a basic tool for the regulation of protein activity (see Section 7.1). Many eucaryotic transcriptional activators are isolated as phosphorylated proteins. The phosphorylation occurs mainly on the Ser and Thr residues, but can also be observed on the Tyr residues. The extent of phosphorylation is regulated via specific protein kinases and protein phosphatases, each being components of signal transduction pathways (see Chapter 7). The phosphorylation of transcriptional activators often represents the final event of a signal transduction chain targeted for a change in gene expression.

An example of how protein phosphorylation can influence the transcription process is the transition from the initiation to the elongation process for RNA polymerase II (see Section 1.4.3.4).

Phosphorylation of transcriptional activators can influence the transcription activity according to the following mechanisms:

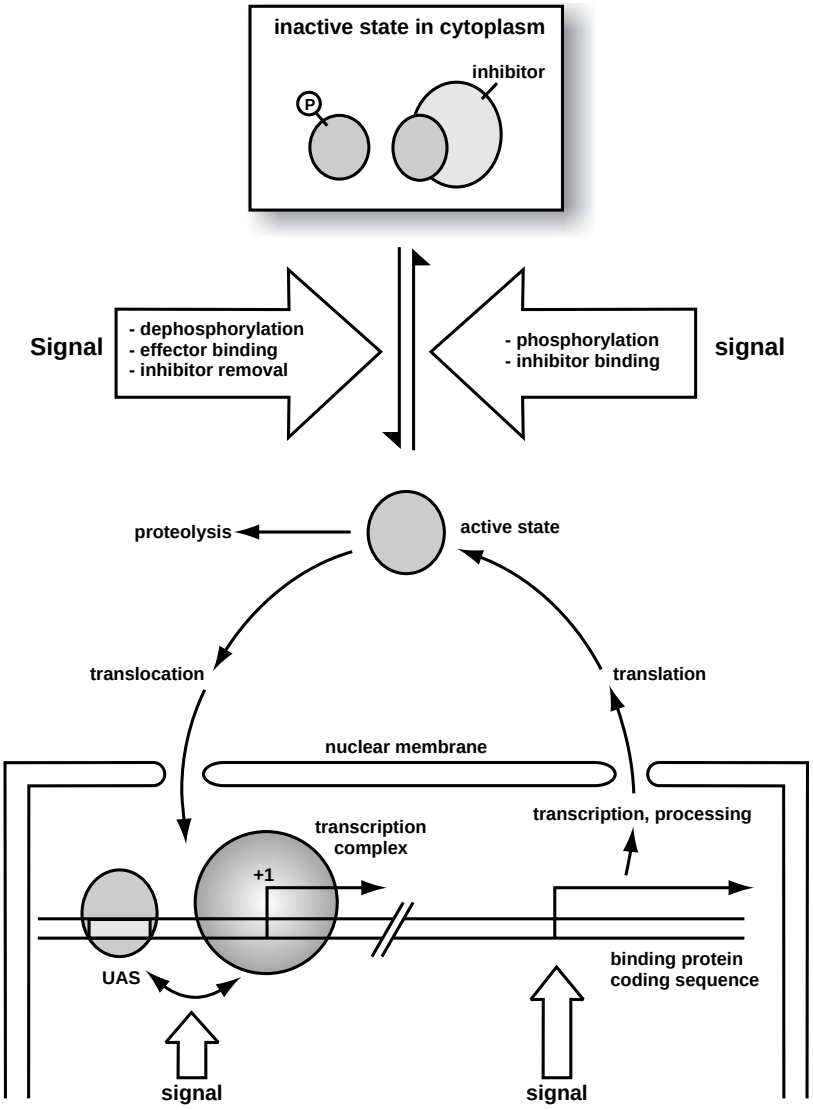


Fig. 1.27 Mechanisms for the control of the activity of transcription factors. Regulatory DNA binding proteins can occur in active and inactive forms. The transition between the two forms is primarily controlled by the mechanisms indicated. Activation or inactivation of transcription factors is determined by signals that become effective either in the cytoplasm or in the nucleus. Signal-directed translocation of transcription factors into the nucleus is a major mechanism for transcriptional regulation. The amount of available transcription factor can also be regulated via its degradation rate or rate of expression. Furthermore, the interaction between DNA-bound activators and the transcription complex can be regulated by various signals.

Regulation of the Nuclear Localization by Phosphorylation

Proteins which act in the nucleus require specific sequences, known as nuclear localization sequences, to direct their transport from the cytoplasm to the nucleus. The *nuclear localization sequences* are generally found at the C-terminus of a protein and often comprise basic amino acids. Phosphorylation in sequences that are required for import into or export from the nucleus can decide whether the transcriptional activator is located predominantly in the cytoplasm or in the nucleus where it can exert its activating function.

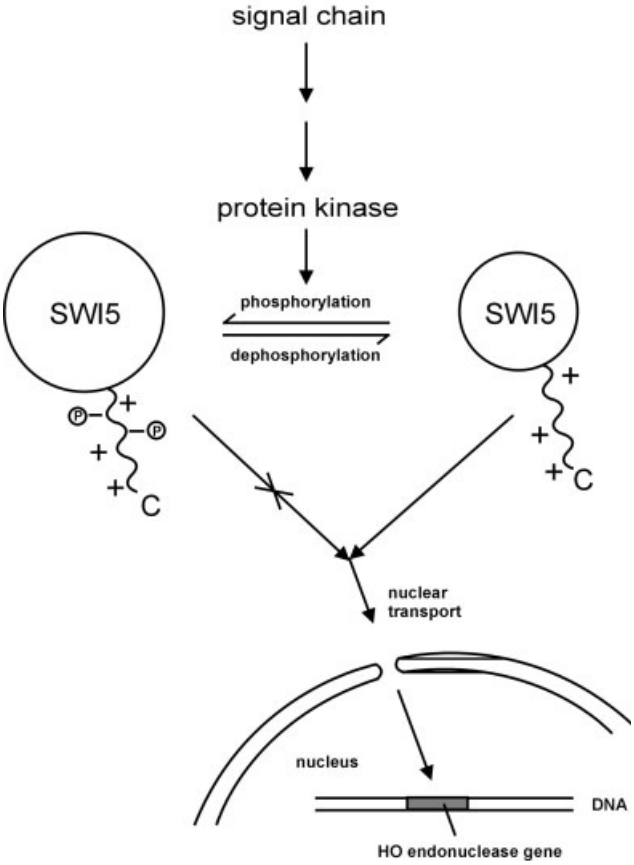


Fig. 1.28 Regulation of the subcellular localization of the transcription factor SWI5 in yeast by phosphorylation. The subcellular localization of the SWI5 protein is regulated by phosphorylation/dephosphorylation. In the phosphorylated state, SWI5 is found in the cytoplasm, while in the under-phosphorylated state it is localized in the nucleus. Phosphorylation and dephosphorylation are catalyzed by either protein kinases or protein phosphatases and can be controlled via signal transduction chains.

An example of this type of regulation is the SWI5 protein of yeast (Fig. 1.28). SWI5 is a transcriptional activator which up-regulates the expression of the HO endonuclease in yeast. SWI5 occurs in two different forms during the cell cycle:

– In the G_1-phase SWI5 is localized in the nucleus and induces the gene for the HO endonuclease.
– In the S-, G_2- and M-phases SWI5 is localized in the cytoplasm and can thus not be active as a transcriptional activator.

The reason for the change in subcellular localization of SWI5 is phosphorylation in the region of the nuclear localization sequence. SWI5 possesses three sequences in the nuclear localization signal for phosphorylation at Ser and Thr residues. Cytoplasmically localized SWI5 is phosphorylated at these positions, thus blocking transport into the nucleus. The protein thus remains in the cytosol.

SWI5 is dephosphorylated at the beginning of anaphase, whereupon transport into the nucleus, binding to the cognate DNA element, and stimulation of transcription become possible. The significance of Ser-phosphorylation for the function of SWI5 has been well documented experimentally. Mutation of the specific Ser residues to nonphosphorylatable Ala leads to a constitutive nuclear localization of the mutated protein and permanent activation of the SWI5-target genes.

The phosphorylation state of the transcription factor NF-AT has a similar effect on translocation. The phosphorylated form of this protein is localized in the cytosol and requires dephosphorylation by the protein phosphatase calcineurin in order to be translocated to the nucleus (see also Section 7.6.5). Other examples of phosphorylation-dependent nuclear translocation include the STAT-proteins (see Section 11.2.2) and the SMAD-proteins (see Section 12.1.2).

In these cases, however, phosphorylation of the transcription factors is required before translocation into the nucleus can occur.

The biochemical basis for the phosphorylation-dependent cytoplasmic localization of transcription factors appears to be a specific interaction of the phosphorylated forms with the nuclear export machinery, allowing the specific export into the cytoplasm of the phosphorylated form only (Kaffman et al., 1998). The preferential export of the phosphorylated form will lead to an increased cytoplasmic localization of the protein.

Phosphorylation of the DNA-binding Domain

There are many examples of the specific phosphorylation of gene-regulating proteins within their DNA-binding domain. Phosphorylation can influence either positively or negatively the ability to specifically bind DNA. The situation becomes complicated by the fact that many transcriptional activators possess multiple phosphorylation sites, whose phosphorylation can have different effects on DNA binding.

How phosphorylation interferes with DNA binding is not well understood. Several mechanisms are conceivable:

– direct interference with the DNA binding due to electrostatic effects
– inhibition of the dimerization of the transcriptional activators
– induction of conformational changes of the protein which cause inhibition or enhancement of DNA binding

Phosphorylation of the Trans-activating Domain

The *trans*-activating domains of transcriptional activators are also common substrates for phosphorylation by protein kinases. The detailed mechanism by which the phosphorylation affects the interactions with the basal transcription apparatus is known in very few cases. The reason for this is the difficulty of identifying the specific interaction partner in the complex transcription apparatus.

Exemplary is the regulation of the CREB protein of higher eucaryotes, displayed in Fig. 1.29.

The CREB protein is a transcriptional activator for genes with *cis*-regulatory, cAMP-sensitive DNA elements (cAMP-responsive elements, CREs).

CREs are DNA sequences which mediate cAMP-regulated transcription. An increase in the cAMP concentration due to hormonal stimulation (see Chapters 5 and 6) activates protein kinases, which can then lead, either directly or indirectly, to phosphorylation and regulation of transcriptional activators. The transcription stimulation of the cognate genes requires the binding of CREB to the CREs and the phosphorylation of CREB at Ser133. This phosphorylation event is mediated by a cAMP-dependent signal transduction pathway. Transactivation by CREB requires a second protein termed CBP (CREB binding protein), which binds specifically to CREB and has the function of a transcriptional coactivator. CBP and a close relative, p300 (review: Chan and Thangue, 2001), have histone acetylase activity and are found as part of a multiprotein complex of varying composition (see Section 1.4.7). The interaction of CBP with CREB depends upon whether it is phosphorylated at Ser133: only if Ser133 of CREB is phosphorylated can CREB and CBP interact. It is assumed that CREB-bound CBP acetylates histones and possibly other chromatin components, thereby relieving a repressed state of the chromatin and allowing the formation of a transcription initiation complex at the transcription start site.

1.4.5.3 Heterotypic Dimerization

Most transcription activators bind to DNA as a dimer or higher multimer (see Section 1.2.4). The dimerization relies on structural motifs which commonly occur in many different proteins. Examples for dimerization motifs are the *helix-loop-helix motif* and the *leucine zipper*. The dimerization motifs permit the formation of DNA-bound *homodimers* or *heterodimers*, depending upon whether the same or different proteins interact with each other (Fig. 1.30). The different dimers may have different requirements for the sequence of the DNA-binding elements, and they can influence transcription activity in very different ways. As shown in Fig. 1.31 for the AP1 family of transcription factors, families of interacting transcriptional activators can be distinguished. AP1 (activating protein 1) is a collective term referring to dimeric transcription activators composed of Jun, Fos or ATF (activating transcription factor) subunits (review: Karin et al.,1997; Shaulian and Karin, 2001), which bind to DNA elements containing closely related recognition sequences. The members of the AP 1 family form homo-or heterodimers that have specific functions in gene regulation. Jun-Jun and Jun-Fos dimers bind preferentially to DNA elements containing the sequence TGACTCA, whereas the Jun-ATF dimers or ATF homodimers prefer to bind to elements containing the sequence TGACGTCA.

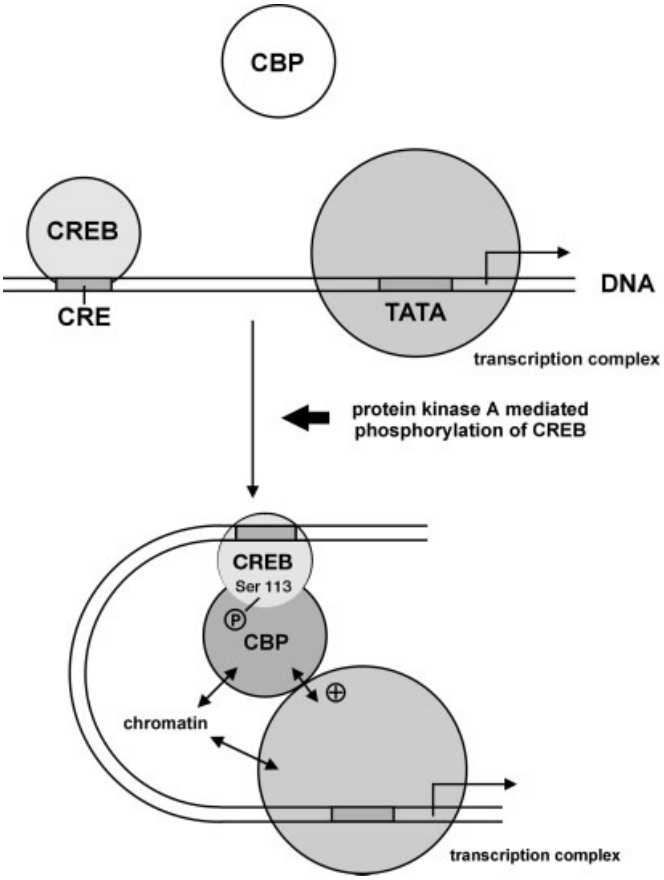


Fig. 1.29 Regulation of the activity of a transcription factor by phosphorylation. The CREB protein is a transcription factor that binds to CREs (CRE: cAMP responsive element) and thereby activates the cognate genes. The CREB protein (CREB: CRE-binding protein) requires the assistance of the coactivator CBP (CBP: CREB binding protein) for efficient transcription activation. The coactivator function of CBP is based on its histone acetylase activity which promotes a transcription-proficient state of the chromatin. CBP can only act as a coactivator if the CREB protein is phosphorylated on Ser113. The phosphorylation of CREB is controlled by a signaling pathway involving cAMP as an intracellular messenger. The DNA element is termed CRE because the cognate gene is regulated by a cAMP-dependent signaling pathway.

Heterotypic dimerization significantly expands the repertoire for tissue-specific regulation of transcription activity. The tissue-specific expression of a particular pattern of transcriptional activators can be used to select only certain DNA-binding elements out of a series of similar elements, and thus to specifically induce certain genes. This strategy is extensively used by the receptors for retinoic acid (see Chapter 4).

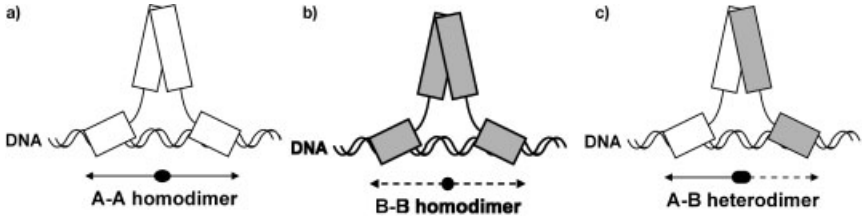


Fig. 1.30 Formation of homo- and heterodimeric transcription factors and the specificy of DNA binding. Shown are two different helix-loop-helix proteins, which bind as homodimers (a, b) to the each of their cognate palindromic DNA elements (drawn as arrows). The two homodimers display different DNA binding specificity. The heterodimerization (c) of the two proteins creates a complex that recognizes a hybrid DNA element.

1.4.5.4 Regulation by Binding of Effector Molecules

The activity of eucaryotic transcriptional activators can be regulated by the binding of low-molecular-weight effectors as well as by the binding of inhibitor proteins (see Section 1.3.3). The most significant example of transcriptional activators regulated by low-molecular-weight effectors is the nuclear receptors, which will be discussed in more detail in Chapter 4.

In this system, cognate hormones act as positive regulating effectors. The transcription regulating activity of nuclear receptors can also be negatively influenced by specific inhibitor proteins. These proteins are characterized as repressors.

1.4.6 Specific Repression of Transcription

A further significant mechanism of transcription control is the repression of gene expression. There are two levels of gene repression to be distinguished in eucaryotes. On the one hand, the chromatin structure can cause an unspecific repression of gene expression (see Section 1.4.7). There is a strong correlation between histone deacetylation, the formation of higher-order structure, and repression of transcription. On the other hand, analogous to the transcriptional activators, many proteins have been characterized that function as repressors of transcription. As in transcriptional activation, repression can be either gene-specific or general. DNA sequences that mediate repression of transcription factors are termed upstream repressing sequences or *silencers*.

Many general repressors function via components of the basal transcription machinery, with the TATA box-binding protein TBP as the major target. Repressors like NC2 are known that bind to the TBP on the promotor and can prevent RNA polymerase II holoenzyme from assembling into the initiation complex. In this way, a general repression of class II genes can be achieved.

The mechanisms of gene-specific repressors are also diverse. Gene-specific repressors may function by binding to transcription activators or by competing with the activator for overlapping binding sites. The extent of repression is then determined

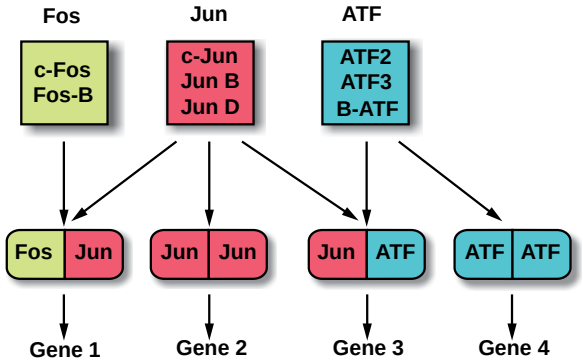


Fig. 1.31 Examples for families of interacting transcription factors. The squares indicate groups of eucaryotic transcription factors, that can form homo- and heterodimers amongst themselves. The members of the Jun family can form complexes with members of the Fos family and with the members of the ATF family. The members of the ATF family can homodimerize or form heterodimers with the members of the Jun family. The Fos family is unique in that its members can not form homodimers, but must heterodimerize with members of the Jun family. The homo- and heterodimers have preference for a distinct DNA element and activate specific sets of genes.

by the relative affinity of both proteins to the DNA element and their concentration ratios.

A further possibility for repression results from heterodimerization (see Section 1.4.5.3). Heterodimers between two transcription factors, in which one of the partners possesses a DNA-binding domain with weak affinity, can inactivate a transcriptional activator in a heterodimer complex. Since a strong binding to the DNA element usually requires both subunits of a DNA-binding protein, transcription activation by this type of heterodimer is not possible.

Gene-specific repressors, analogous to transcriptional activators, are often constructed modularly, with a DNA-binding domain and a repressor domain. The repressive character of such domains has been proven in domain-swapping experiments. Other gene-specific repressors are recruited to specific promotors by interaction with a DNA-binding protein and function by forming repressive contacts to the transcription apparatus or by promoting a transcription-incompetent structure of chromatin.

An interesting and functionally important aspect of transcriptional activation and repression is that one and the same protein can act as both an activator and a repressor. The alternative functionality is determined by the sequence environment, by specific repressors, or by low-molecular-weight effectors. Examples are the receptors for vitamin A acid, which, in the presence of its ligand, vitamin A acid, activate the genes with cognate DNA elements. In the absence of the ligand, the same receptor has a corepressor associated and represses transcription in the DNA-bound form (see Chapter 4). The corepressor is an integral part of a larger protein complex that binds the

vitamin A receptor and appears to weaken the transcription of the gene via the activation or recruitment of a histone deacetylase activity (see Section 1.4.7). It is assumed that the repressor protein dissociates from the receptor in the presence of the cognate hormone to enable transcription activation to occur.

In summary the following mechanisms of general and specific repression have emerged (see Fig. 1.32):

- Direct inhibition of the formation of a pre-initiation complex: complexation of basal transcription factors, such as TBP or TFIIB, or competition with TFIIB for binding

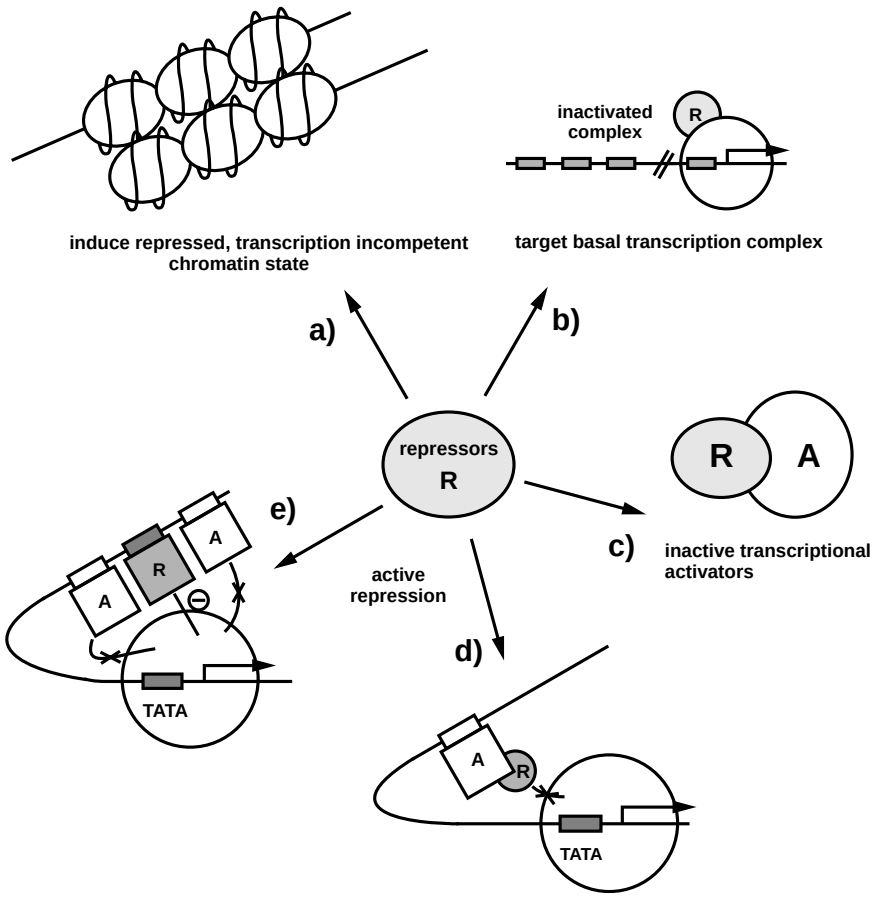


Fig. 1.32 Pathways of repression of transcription. The figure illustrates various mechanisms of repression of transcription. a) Repressors can induce a generally repressed state in chromatin which is incompatible with transcription. To allow transcription at all, the repressed state must be relieved. b) Repressors can target the transcription complex and thereby inhibit transcription initiation. c, d) By binding to free or DNA-bound transcriptional activators, repressors can block the activating function of the latter. e) active repression is also affected by proteins that bind sequence specifically to DNA elements and in their DNA-bound form inhibit the transcription initiation.

to the promotor. An example of this type of repression is the negative cofactor NC2 (see Section 1.4.3). Transcription repression can also result from phosphorylation of the basal transcription factors. By this token, the repression of transcription observed during mitosis (see Chapter 13) is attributed to the hyperphosphorylation of TBP and TAFs.

- Inactivation of the *trans*-activating protein by specific complex formation. Examples include the Gal80 inhibitor from yeast that complexes and inactivates the Gal4 activator .
- Inhibition of the transition from the initiation phase to the elongation phase.
- Induction of a chromatin structure that does not allow the efficient formation of the pre-initiation complex, e.g., by deacetylation or by methylation of histones (see Section 1.4.7).

1.4.7 Chromatin Structure and Transcription Activation

In the previous considerations, the function of chromatin structure on transcription activation has been ignored. Generally it holds true that chromatin structure is decisive for gene activity, and certain configurations of chromatin are associated with transcription repression.

Chromatin Structure

The basic structural unit of chromatin is the nucleosome, in which 146 bp of DNA are wrapped 1.65 turns around the histone octamer (H2A, H2B, H3, H4)$_2$. The co-crystal structure, a histone-octamer-DNA complex, shows that the DNA wraps tightly around a cylinder-like core of the histones. The DNA is arranged such that interactions between the N-terminal tails of histones and adjacent nucleosomes are possible. Furthermore, the DNA wrapping around the histone core is irregular, providing some flexibility in the structure, which may be relevant for remodeling of the nucleosome during transcription.

The nucleosomes are further packaged into higher-order structures, among which the so-called solenoid is best characterized. Linker histones like histone H1 are involved in this organization. Formation of the higher-order structures generally has a repressive influence on transcription.

Covalent Modifications of Histones

Histones are subject to the folowing covalent modifications (Fig. 1.33; review: Lee and Young, 2000):

– acetylation, mostly of lysine residues
– phosphorylation at Ser/Thr residues, e.g., of histone H3
– methylation at lysine residues
– ubiquitination.

Often these modifications are found at the N-terminal tails of the histones, and their presence correlates with a transcriptionally active state of the chromatin. The multiple modifications of histone tails may serve as a combinatorial code for instructing cellular actions on the DNA template during mitosis, transcription or replication. Of the different modifications, the acetylation of histones is of outstanding importance for transcription regulation and will be discussed later (see Section 1.4.7.1). Furthermore, histone methylation has now been recognized as another important regulatory tool for transcriptional control (see Section 1.4.7.2).

The phosphorylation of histone H3 can have a repressive or activating effect on transcription. Members of the MAP kinases (see Chapter 10) that are responsible for this phosphorylation have been identified, providing a link between growth factor stimulation and transcription activation at the level of histone modification. H3 phosphorylation has been also recognized as part of a complex signaling mechanism that operates in the condensation/decondensation of chromatin during the cell cycle. Furthermore, histone H1 phosphorylation has been linked to the relief of transcription repression.

The targeted ubiquitination and subsequent degradation (see Chapter 2) of histones also seems to have a specific role in regulating transcription via chromatin modification. One member of the TAF proteins, namely $TAF_{II}250$, has been shown to mediate the ubiquitination and subsequent proteolytic degradation of histone H1. Since the presence of histone H1 generally has a repressive effect on transcription, this mechanism may help to alleviate repression of transcription by H1.

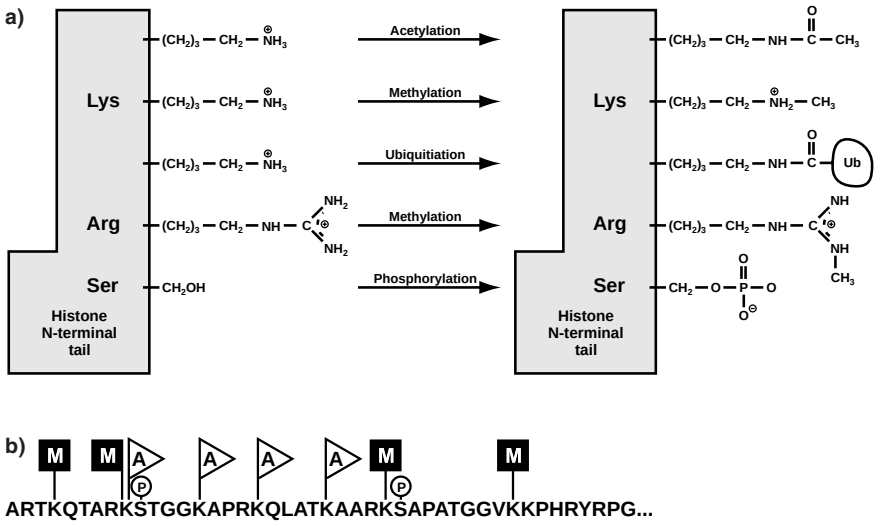


Fig. 1.33 a) Covalent modifications of histones b) Modifications of histone H3. The amino acid sequence of the N-terminal tail of histone H3 from mammals is shown with some of the covalent modifications. Modification reactions are influenced by the prior modification state as indicated for some H3 cases. A: acetylation; M: methylation, P: phosphorylation.

Chromatin Remodeling

Nucleosomes are subjected to conformational changes during transcription initiation and elongation. This noncovalent modification results in a rearrangement of histone-DNA contacts and a mobilization of the nucleosomes on the DNA template (review: Narlikar et al., 2002). Several protein complexes have been shown to be involved in chromatin remodeling. Examples are the SWI/SNF complex from yeast and human and the human ACF and NuRD complexes. A common characteristic of the remodeling complexes is the presence of distinct ATPase subunits. The ATP-dependent remodeling complexes use ATP hydrolysis to increase the accessability and mobility of DNA sequences bound to the nucleosomes. They catalyze fluidity in the position and conformation of the nucleosomes, resulting in, e.g., a specific positioning of nucleosomes with respect to the transcription start site. It is assumed that the energy of ATP hydrolysis is used to disrupt nucleosome-DNA contacts and favor the transition between various conformational states of nucleosomes and chromatin.

Transcription of Nucleosome-covered DNA Templates

The presence of chromatin represents a block to transcription. To relieve this block, a modification and remodeling of the nucleosomes and of the whole chromatin is required both for the initiation and elongation steps. In actively transcribing genes, the nucleosomes are not displaced from the DNA template. Rather, they are mobilized and can be shifted during initiation and elongation of transcription to downstream sites. Furthermore, RNA polymerase II is able to transcribe through nucleosome-covered DNA quite efficiently in the presence of chromatin remodeling complexes. In this process the nucleosomes are modified as outlined above whereby the proteins that are necessary for the restructuring of the nuceosomes are carried along with the transcribing RNA polymerase.

A *defined rearrangement of nucleosomes* has been clearly shown.during activation of the interferon-ß (IFN-ß) gene (Agliatoli et al., 2000; Lomvardas and Thanos, 2001). Here, the blocking action of nucleosomes bound to the enhancer region and to the core promotor is relieved by the combined action of SWI/SNF and HAT complexes. The nucleosomes present at these sites are shifted to neighbouring sequences providing space for deposition of the enhanceosome and the basal transcription apparatus.

Interestingly, nucleosomes can potentiate as well as repress gene activation. The stimulatory effect of nucleosome positioning appears to be related to the specific exposure of recognition sequences for transcription activators. In addition, specific nucleosome positioning may help to create a specific DNA topology which may facilitate (or hinder) transcription factor binding.

Thus DNA packaging into chromation appears to provide a distinct physical context for the protein-DNA and protein-protein interactions that are required for a specific and efficient initiation.

It is becoming increasingly evident that multiple linkages exist between transcription and chromatin. A large part of the function of transcription factors is dependent on and directed to the structure of chromatin. Multiple linkages exist between chromatin and transcription. Within this network, interactions between transcriptional

activators (or repressors) and the following multiprotein complexes have been shown to occur:
– RNA polymerase II holoenzyme
– chromatin-remodeling machines
– chromatin assembly factors
– histone acetylase and histone deacetylase complexes
– histone methyltransferases.

A model of the action of chromatin-modifying proteins during transcription initiation and elongation is shown in Fig.1.34. An initial disruption of chromatin is achieved by transcriptional activators that upon binding to their cognate DNA element target chromatin-modifying proteins to the transcription start site. The DNA elements of the activators may be accessible also on nucleosome-covered DNA. Modification and repositioning of the nucleosomes in the vicinity of the DNA-bound activators will give space for the binding of RNA polymerase II holoenzyme to the promotor. The phosphorylation of the CTD of RNA polymerase II that signals the transition from initiation to elongation is thought to be also a signal for the recruitment of chromatin-modifying activities like PCAF into the RNA polymerase holoenzyme complex. The chromatin modifiers than travel along with RNA polymerase II, thereby propagating the disruption and modification of chromatin.

Among the different modifications of chromatin, the acetylation/deacetylation of histones has been studied best and will therefore be discussed separately in the following section.

14.7.1 Transcriptional Activity and Histone Acetylation

Numerous studies have shown that the acetylation or deacetylation of the histones of the nucleosome plays a major role in the regulation of transcriptional activity. Acetylation of the histones (review: Narlikar et al., 2002) is a post-translational modification, which is usually performed on lysine residues at the N-terminus and requires specific enzymes, the *histone acetyl transferases* (*HATs*). Removal of the acetyl group also requires specific enzymes, the *histone deacetylases* (*HDAC*). Most importantly, the acetylation of histones is accompanied by a loss of positive charges, which is thought to have a profound influence on the nucleosome structure and on the strength of DNA binding.

Histone Acetylases

A large number of HATs and HAT-complexes have now been described (reviews: Marmorstein, 2001; Goll and Bestor, 2002). The purification and biochemical characterization of the HATs has shown that these enzymes also can acetylate substrates other than histones, includig transcription activators such as p53 and GATA-1, structural proteins like HMG1, and the general transcription factors TFIIE and TFIIF. In addition, the acetylation of histone acetylases themselves has been observed.

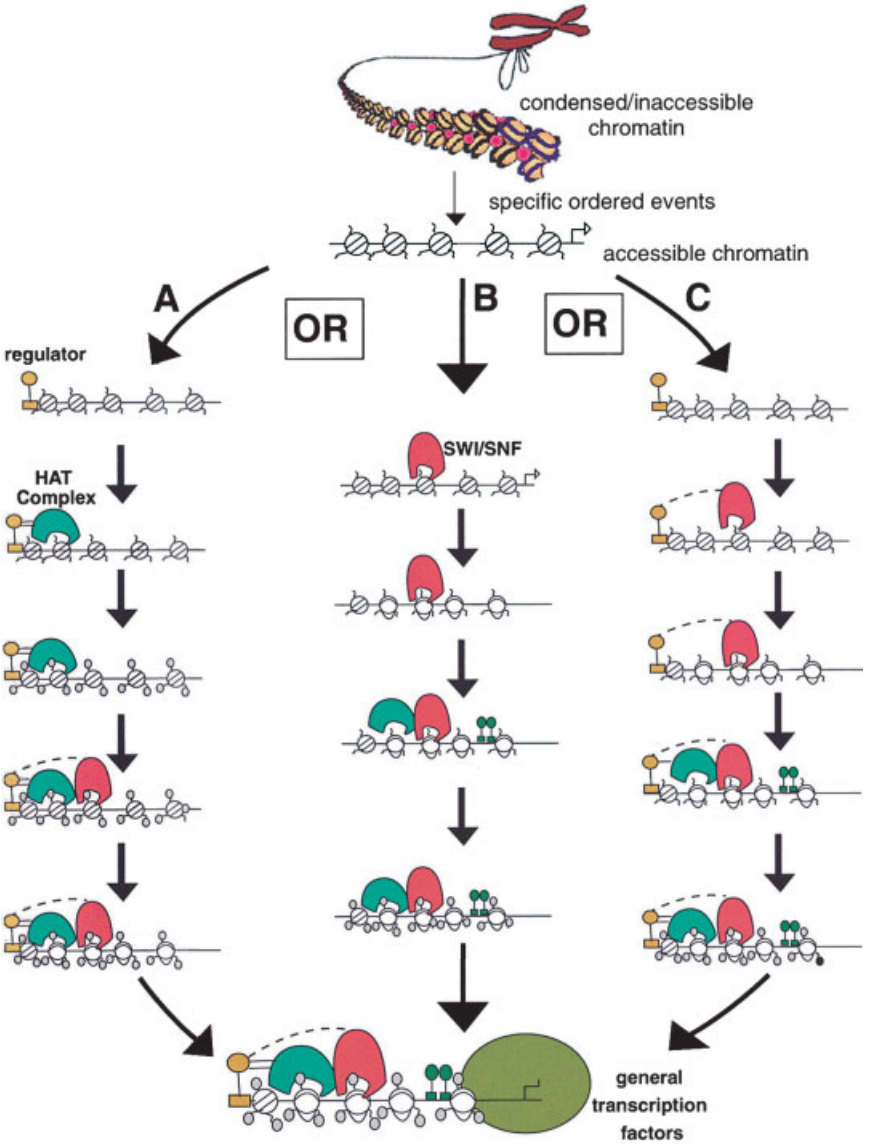


Fig. 1.34 Model of the action of transcriptional regulators and chromatin-remodeling complexes (After Narlikar et al., 2002). To prepare the DNA template for transcription, condensed chromatin is reorganized and becomes accessible for the binding of transcriptional regulators and remodeling complexes. In the model shown, the transcriptional regulator binds first to its DNA element and then recruits histone acetylase (HAT) complexes to the region of accessible chromatin. In the next step, the remodeling complex SWI/SNF binds to this region leading to reorganization and increased fluidity of the nucleosomes. The basal trancription apparatus can then bind to the promoter and the gene section is now competent for transcription. Only one possible order of events is shown in the figure. Other pathways, where the sequence of events is different, are equally possible (see Narlikar et al., 2002).

Wheres the consequences of acetylation of these other substrates remain ill understood, the function of acteylation of histones clearly is in transcriptional activation. In many, although not all, cases, the histone acetylation is correlated with a stimulation of transcriptional activity. HAT activity is found in proteins associated with the transcription apparatus or identified as coactivators. Examples are, among others, members of the TAF proteins including $TAF_{II}250$, the GCN5 protein (yeast and human) and the coactivator CBP, and the related protein, p300. Generally, the histone acetylases are part of large protein aggregates of complex and varying composition (review: Struhl and Moqtaderi, 1998), where one and the same HAT enzyme can occur in different complexes. Several types of HAT-containing complexes have been characterized: the SAGA and the ADA complexes from yeast and the human PCAF. These complexes contain several histone-like TAFs and HAT enzymes, GCN5 in the yeast SAGA and ADA complex, and a human GCN5 homolog in the PCAF complex.

The HAT enzymes can be grouped by cellular localization into two classes. Typ A HATs are localized in the nuclei and acetylate nuclear proteins, while, in contrast, type B HATs are found in the cytoplasm and probably acetylate newly synthesized histones.

A large part of the HAT activity functions in a gene-specific manner via specific interactions between transcription activators and the HAT complexes. We now know of numerous examples of transcription activators that target HAT activity to specific genes. A HAT enzyme recruited by many activators is p300/CBP, which is found as part of the PCAF complex. Examples of transcription activators that recruit HAT activity are

- Swi5 from yeast
- c-jun
- CREB
- nuclear receptors
- Stat factors.

As already mentioned, the HAT enzymes are found in most cases as part of the coactivator complexes required for the function of these activators.

The mechanism of transcriptional activation via histone acetylation remains speculative. Two models, which are not mutually exclusive and may cooperate, are currently under discussion:

- Alteration of nucleosome structure: the neutralization of positive charges of the histones by acetylation leads to a modification, or loosening, of the nucleosome structure (Fig 1.35). The loss of positive charges will weaken the interaction of the nucleosomes with the DNA. As consequence, the mobility of the nucleosome on the DNA will be increased, and the DNA may be more accessible for transcriptional activators.
- Binding of additional accessory proteins: the acetylated histones can serve as a platform for the association of further protein components with the genetic element, so that transcription of the gene is modified.

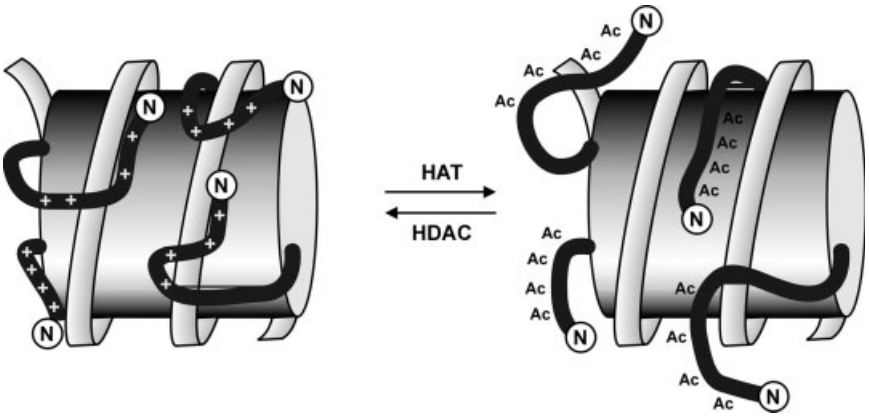


Fig. 1.35 Model of the influence of histone acetylation and deacetylation on the nucleosome structure

The amino terminal tails of the four core histones contain lysines that are acetylated by HATs and deacetylated by HDACs. The histone octamer (H2A, H2B, H3, $H4)_2$ is represented as a cylinder wrapped by DNA. It is thought that neutralization of the positive charges on the histone tails results in alterations of the nucleosome structure and of associated non-histone proteins that may lead to a higher mobility of the nucleosome and/or an improved accessability of the bound DNA with accompanying changes in chromatin structure, chromatin hierarchy and transcription. In most, but not all cases, deacetylation correlates with the repressed state and acetylation correlates with the transcriptionally active state. The amino termini of the histones are indicated by a circled N; charged lysines are represented by +; and acetylated lysines are indicated by Ac. Only the hyper- and hypoacetylated states are depicted.

Histone Deacetylases

The histone deacetylases were first identified as components of large protein complexes with corepressor function in transcription regulation. Numerous HDACs have now been described, indicating a large diversity of the deacetylase enzymes in eucaryotic cells (review: Tong, 2002). In addition to a role in transcription repression, HDACs appear to be involved also in other major functions of cells as, e.g., ageing and long-term silencing.

By opposing the activity of the HATs, the HADCs are thought to stabilize the nucleosome structure and to decrease nucleosome mobility, rendering the DNA inaccessible for binding of the proteins required for transcription activation.

The deacteylase complexes are targeted to specific promotors by interactions with different types of negative regulatory proteins. Among these are sequence-specific DNA-binding proteins, corepressors and the methylated CpG-binding proteins (see Section 1.4.8)

Most evident is the association of HDAC activity with transcriptional repressors or corepressors. As an example, the repressive heterodimeric transcription factor Mad-Max forms a complex with the histone deacetylase HDAC I that is part of the mammalian mSin complex. A complex of HDAC I and the nuclear receptor-corepressor (see Chapter 4) bind to unliganded nuclear receptors and are believed to exercise a repressive effect. A further example is the tumor suppressor protein pRb (see Chap-

ters 13 and 14), which can occur as a transcriptional corepressor in the hypophosphorylated form and a transcriptional coactivator in the hyperphosphorylated form. The repressive form of the pRb protein recruits the histone deacetylase HDAC I to the DNA and thereby initiates an active repression of the gene (see Section 13.4.2).

1.4.7.2 **Transcriptional Activity and Histone Methylation**

Methylation of Lys and Arg residues of histones has been identified as another tool for modification of histone structure and for regulation of transcriptional activity. The methylation of lysine residues of histones is an example of an epigenetic code that is used to establish transcriptionally inactive states both in heterochromatin and in euchromatin.The best characterized example is histone H3, which is found to be methylated at K4, K9 and K20. These methylations have a differential effect on transcriptional activity.

Methylation at K9 is associated with transcriptional inactivation and establishment of a heterochromatin state. K9 methylation by the methyltransferase SUV39 directs binding of the transcriptional repressor HP1, which is a component of transcriptionally inactive heterochromatin (Fig. 1.36). SUV39, HP1 and histone deacetylase enzymes have been found to be associated with inactive, hypophosphorylated complexes between the tumor suppressor pRb and the transcription factor E2F (see Chapter 13).

These observations indicate that the SVU39/HP1 complex is involved in transcriptional silencing at heterochromatic loci and can be specifically recruited to genes by association with transcriptional regulators.

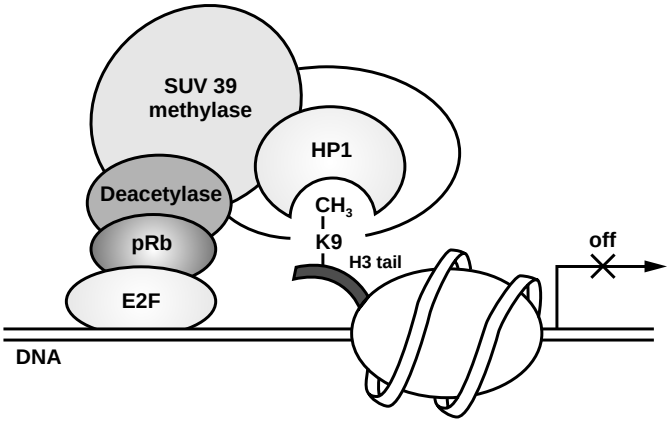


Fig. 1.36 Methylation of H3 during silencing of pRB-E2F-controlled genes
The tumor suppressor pRb interacts with the transcription factor E2F and can repress transcription by recruiting histone deacetylase activity to E2F-controlled genes. The deacetylase removes an acteyl group from Lys9 of H3. In a subsequent reaction, the methylase activity of SUV39 attaches a methyl group to Lys 9 providing a binding site for the repressor HP1. Binding of HP1 then generates a highly compact chromatin structure by a still unknown mechanism. HP1 and the SUV39 methylase are found in tight complex in the cell (After Kouzarides, 2002).

In contrast, methylation at K4 of H3 is correlated with transcriptionally active states of chromatin. The K4 methylation appears to be absent from histones with methylation at K9, but colocalizes with acetylation at K9, indicating a positive cooperation of acetylation and methylation of distinct Lys residues during the establishment of a transcriptionally active state. For yeast histones, a link between methylation and acetylation at Lys residues and phosphorylation at Ser residues of histone H3 has been discovered (Nakayamaet al.,., 2001), defining a conserved pathway of sequential histone modifications during heterochromatin assembly. These modifications are recognized by specific proteins that are part of multiprotein assemblies in heterochromatin.

Overall, a complicated picture of cooperative modifications of histones has emerged. It has been proposed that distinct modification of histone tails is used to specify a "code" that dictates the regulatory features of a cell (review: Berger, 2001). The "code" is then read by a set of proteins linking the modification pattern to either the active or repressed chromatin state. Within a given gene, the code may change from one state to the other, e.g., during progression through the cell cycle. An interplay exists between different modifications in the sense that the presence of a given modification may promote or inhibit modification elsewhere on the same histone molecule.

1.4.7.3 **Enhanceosomes**

Another level of complexity in gene regulation in higher eukaryotes has been recognized to exist in the form of enhanceosomes. These are multi-protein complexes formed at enhancer sequences that register and integrate gene regulatory signals. Typical enhancer sequences are located close to or at a greater distance from the promotor, and function independently of their orientation relative to the transcription start site. Furthermore, enhancers often comprise several DNA elements that serve as binding sites for distinct transcriptional regulators.

Cooperative binding of architectural and regulatory proteins to binding sites of the enhancer creates a multiprotein-DNA complex named an enhanceosome, which is essential for the initial modifications of chromatin at promotors and for stable binding of the core transcription apparatus. Protein-induced DNA bending, histone modification, histone mobilization, and cooperative, combinatorial interactions between different transcription factors are major tools used in enhanceosome-mediated transcriptional regulation.

The best studied example is the interferon-β (IFN-β) enhanceosome (review: Merika and Thanos, 2001) that assembles at the enhancer of the IFN-β gene and is responsive in a highly specific way to virus infection. The IFN-β enhancer contains three overlapping regulatory DNA elements recognized by the transcriptional activators NFκB, the interferon regulatory factor (IRF), and the ATF/c-Jun heterodimer. In its natural context, the IFN-β enhanceosome responds only to the signal triggered by virus infection, whereas synthetic constructs comprising only two of the binding sites can respond to other inducers too, indicating that the identity and regulatory function of the enhanceosome is specified by a cooperative interaction between all three types of transcription factors.

Central to the function of the enhanceosome appears to be the protein HMG I(Y), which is not a typical transcription activator but rather an architectural DNA-binding protein. HMG I(Y) belongs to the family of high-mobility group proteins which comprises proteins with specific or nonspecific DNA-binding properties. HMG I(Y) binds to four sites on the enhancer and contacts simultaneously the transcription factors inducing the cooperative formation of a multiprotein-DNA complex. This complex also recruits histone acetylase activites residing in the PCAF complex or in CBP (see above), allowing for acetylation of histones and of HMG I(Y) itself. In this process, the architectural function of HMG I(Y) is fine-tuned by acetylation in a dynamic way, and a distinct rearrangement of nucleosomes is observed (Fig. 1.37).

One important step appears to be the modification and mobilization of a nucleosome positioned near the TATA box that blocks TBP/TFIID binding in the uninduced

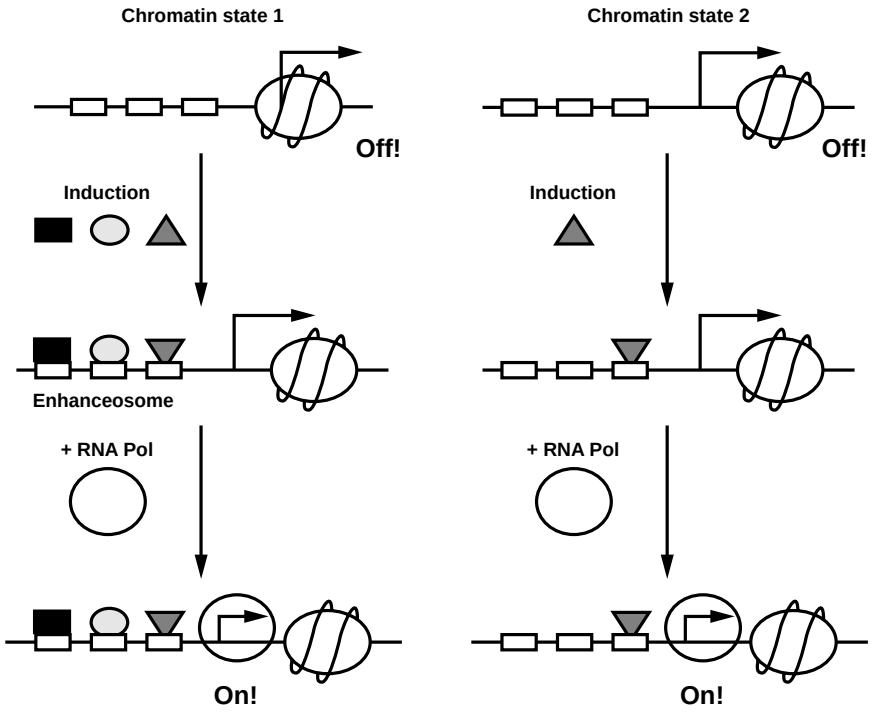


Fig. 1.37 Model of enhanceosome function during transcription activation. The requirement for only a single regulator or multiple transcriptional regulators, organized in the enhanceosome, may depend on the state of chromatin. In chromatin state 1, the promotor is covered by a nucleosome and binding of three regulators (circle, square, triangle) is required to achieve effective recruitment of chromatin modification complexes and to free the promotor for binding of the RNA polymerase holoenzyme. In this case, enhanceosome formation is needed to recruit the chromatin modifying proteins and induce transcription of the gene. Chromatin state 2 depicts a situation where the promotor is not shielded by a nucleosome and is already accessible for the transcription apparatus. In this situation, binding of a single regulator may be sufficient to bring about stable binding of the transcription apparatus to the promotor (After Lomvardos and Thanos, 2002).

state. Upon induction, the enhanceosome recruits histone acetylase activity and SWI/SNF activity leading to acetylation of this nucleosome and shifting to a downstream position. Binding of TBP to the TATA box is now possible and the induced bending of the DNA stabilizes a position of the nucleosome distant from the transcription start site. The core promotor is now fully exposed which is a prerequisite for transcription initiation.

As shown by studies on other enhancer elements, the sequence of events depicted in Fig. 1.37 is not fixed and depends on the nature of the enhancer elements and on the specific chromatin structure at the promotor of the gene. Promotors with less restrictive nucleosome arrangements are thought to require less steps for activation of transcription initiation and may therefore be activated by a distinct set of stimulatory signals. Depending on the local chromatin structure, the transcription activators recruit different protein complexes to the chromatin. If the chromatin structure is inhibitory, then the modifiers/modulators are required to alter it. A decisive point is the ability of the activators to recruit distinct protein complexes to chromatin in response to different activating signals. Which of the complexes is used depends on the nature of the activating signal and on the local chromatin structure which is therefore another layer in transcriptional regulation.

1.4.8 Methylation of DNA

The methylation of DNA is the most abundant epigenetic modification in vertebrate genomes. In the mammalian genome, the methylation takes place only at cytosine bases that are located 5'to a guanosine in a CpG dinucleotide (Fig. 1.38). A high CpG content is found in regions known as CpG islands. Most CpG islands are found in the vicinity of promotors and are unmethylated in normal cells whereas in the remainder of the genome, the CpG sequences are generally methylated. In specific situations as e.g. during tumorigenesis, CpG islands are however found in a hypermethylated state (see Section 14.1.3).

A characteristic distribution pattern of 5-methyl cytidine (m^5C) is found within each cell, which remains intact upon cell division. Mechanisms must therefore exist to ensure that the methylation pattern is precisely retained in the daughter cells following cell division. The enzymes responsible for DNA methylation at CpG sequences are the DNA methyltransferases. The methyl group is derived from S-adenosyl methionine. Several DNA methyltransferases have been characterized, some of which, e.g., Dnmt1, can perform *hemi-methylation* in the CpG sequences (Fig. 1.38). The preferential substrates for hemi-methylation are DNA sequences in which the complementary strand is already methylated. Such a hemi-methylation occurs, for example, shortly after replication of the sequence. This type of DNA methylation is also called maintenance methylation and is responsible for the inheritance of the methylation pattern. In addition to maintenance methylation, *de novo* methylation at CpG sequences is also possible. An important inhibitor of DNA methyltransferases is 5-aza-cytidine, which blocks DNA methylation leading to a change in DNA methylation patterns of cells.

DNA Methylation and Gene Repression

A strong correlation exists between the methylation status at CpG sequences and the transcription activity:

- Constitutively inactive genes display a high density of methylation; active genes tend to be undermethylated.
- Transcriptional activity can be modulated by the methylation of CpG sequences near promotors.
- Inhibition of methyl transferase by 5-azacytidine (Fig. 1.38) leads to a change in the methylation pattern and to a reversal of differentiation of a cell culture.
- Methylation of foreign DNA, such as viral DNA, provides a defense mechanism against the expression of exogeneous DNA.
- DNA methylation participates in genetic imprinting and in X-chromosome inactivation. The term "genetic imprinting" describes a situation where genes are expressed unequally depending upon whether they were maternally or paternally inherited. Normally both copies of the parental genes are equally transcribed in a diploid chromosome. However, with imprinting, a gene inherited from either the mother or father is selectively inactivated. Methylation is obviously involved in such an inactivation. The inactive copy is more strongly methylated than the active copy.
- In specific situations of cells, e.g., during cell cycle progression and differentiation, the methylation pattern can be quite dynamic. Enzymes have been identified that function as demethylases and can alter the methylation pattern. Furthermore, DNA methylation has been recognized as an important aspect of tumorigenesis. Changes in the methylation pattern have been linked in many tumors to decreased expression of tumor suppressor genes (see Chapter 14).

Mechanism of Gene Repression by DNA Methylation

Mechanistically, the repressive function of CpG methylation has been linked to methyl-CpG-binding proteins that specifically recognize and bind to methylated CpG sequences. The proteins MeCP2, MBD1, MBD2, MBD3 and MBD4 constitute a family of vertebrate proteins that share a domain called the methyl-binding domain (MBD). MeCp2 and MBD1-3 were found as components of large corepressor complexes and have been shown to direct histone-deacetylase activity to chromatin (Nan et al., 1999; review: Ballestar and Wolffe, 2001). By recruiting histone deacetylase activity to methylated CpG sequences, a repressed state of chromatin is established in the vicinity of the CpG sequences. The nucleosomes and other chromatin components are under-acetylated under these conditions, and the chromatin is inaccessible for the transcription machinery.

In summary, we now know of three types of epigenetic modifications that are used to maintain a repressed stae of chromatin and to control the flux of information from DNA to RNA (review: Richards and Elgin, 2002). These modifications are (Fig. 1.39):

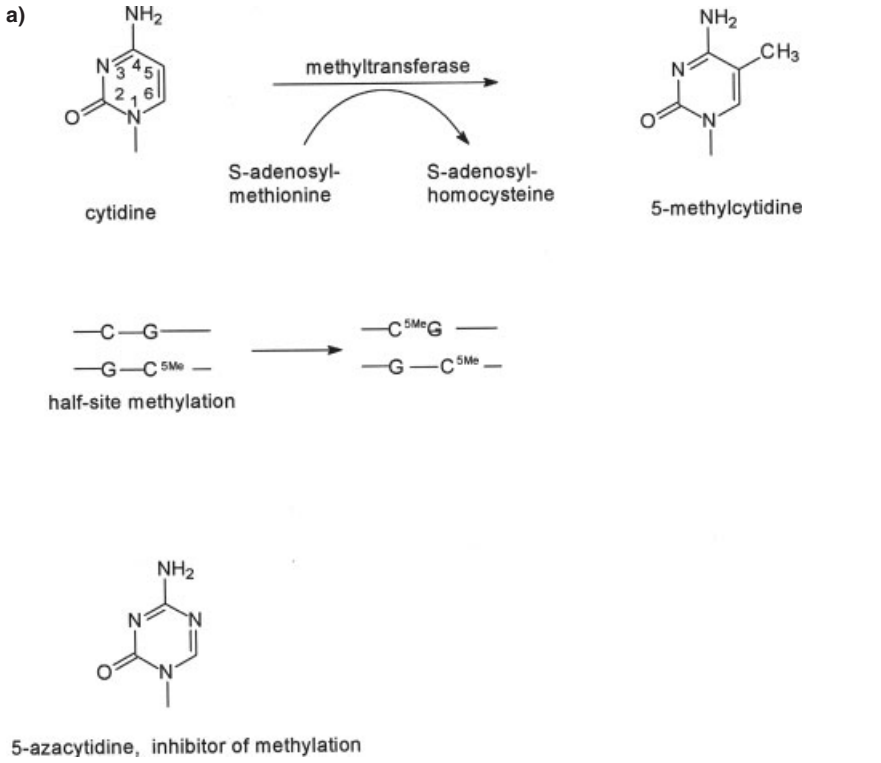


Fig. 1.38 The methylation of DNA: 5-methyl-cytidine and maintenance methylation. a) The methylation of cytidine residues on DNA is catalyzed by a methyl transferase that employs S-adenosine methionine as a methyl group donor. The peferable substrate for the methyl transferase are hemi-methylated CpG sequences. 5-aza-cytidine is a specific inhibitor of methyl transferses. b) The methylation pattern of DNA remains intact upon DNA replication and is passed on to the daughter cells. The newly synthesized strands are unmethylated immediately after DNA replication. The methyltransferase uses the previously methylated parent strand as a matrix to methylate the CpG sequences of the newly synthesized strand.

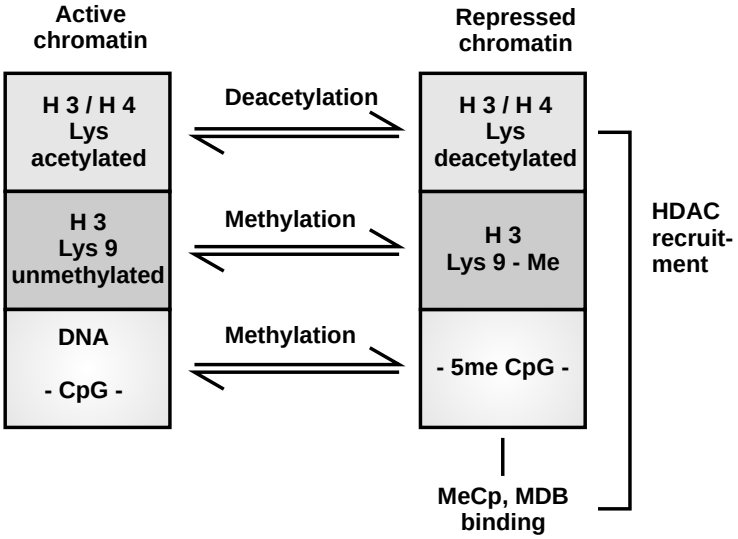


Fig. 1.39 Epigenetic modifications of chromatin
Transition of active to inactive states of chromatin (and vice versa) is controlled by epigentic modifications including acteylation of N-terminal Lys residues of histones H3 and H4, methylation of Lys9 of H3 and methylation of CpG sequences. CpG methylation induces binding of Methy-C-binding proteins (MeCP and MDB proteins) which in turn leads to recruitment of histone deacetylase activity (HDAC) establishing a repressed state of the chromatin.

- Histone hypoacteylation: The recruitment of HDACs to chromatin stabilizes transcription-incompetent states of nucleosomes.
- Methylation of Lysine 9 of Histone H3: This modification directs proteins like HP1 to chromatin that maintain a repressed state.
- Methylation of DNA at CpG sequences. Methylated CpG elements are recognized by methyl-CpG binding proteins that recruit histone deacetylase activity and thereby induce an inhibitory chromatin configuration.

It is assumed that a self-reinforcing network of interactions exists between these modifications that is used in packaging heterochromatin and in stable silencing of euchromatic genes.

1.5 Post-transcriptional Regulation of Gene Expression

Transcription and translation are spatially separated events in eucaryotes. The product of nuclear transcription is *pre-mRNA*. In order to enable translation, the information contained within the pre-mRNA must be transported out of the nucleus and into the

cytosol. The quantity of processed mRNA available for translation decides to a high degree how much protein is formed by *de novo* synthesis.

From the primary transcript to translated protein there are many possible points for regulatory processes. The most important regulatory points are
– modification of the 3' end of the pre-mRNA
– splicing of the pre-mRNA
– transport of the pre-mRNA
– initiation of translation
– stability of the mRNA.

Modification at the 5' and 3' end of the pre-RNA, as well as splicing of the primary transcript are the major modifications that are necessary to form the mature mRNA ready for translation at the ribosome. The 3' end modifications and splicing decide which information contained in the primary transcript is made available for protein biosynthesis. The information content of the processed mRNA can be specifically influenced by these processes. This has an important impact on the tissue- and cell-specific protein expression.

3' and 5' modifications and splicing are tightly coupled to transcription by RNA polymerase II. Interventions in these processes are another possibility for regulation at a level that is distinct from transcription *per se*. The intimate connection between transcription and pre-mRNA processing has been recognized only in recent years, when it became clear that components of the mRMA-modification systems are associated with the Carboy-terminal domain (CTD) of elongating RNA polymerase II.

The translation of the correctly modified mature mRNA by the ribosome is also subject to regulation. The regulatory site of translation is mainly at the initiation of translation. Further regulatory elements include the availability of mRNA for ribosomal protein biosynthesis, as well as the concentration of mRNA. The availability of mRNA can be controlled by, for example, sequence-specific protein binding to the mRNA. The concentration of a specific mRNA is determined by a balance between its rate of synthesis (i.e. transcription) and its rate of degradation by RNases. The stability of an mRNA against nucleolytic degradation is thus a further factor that can determine the extent of biosynthesis of a protein.

1.5.1
Modifications at the 5' and 3' Ends of the Pre-mRNA

Modifications at the 5' and 3' ends include the processes of capping and polyadenylation (Fig. 1.40).

Capping at the 5' end of the pre-mRNA occurs immediately after incorporation of about 30 nucleotides in the primary transcript. The enzymes involved in capping become associated with the CTD of RNA polymerase II immediately after CTD hyperphosphorylation, at the transition from initiation to elongation. The 5' cap structure is required for the binding of the mRNA to the 40S subunit of the ribosome during the initiation of translation. Also, a stabilizing function for mRNA is ascribed to capping.

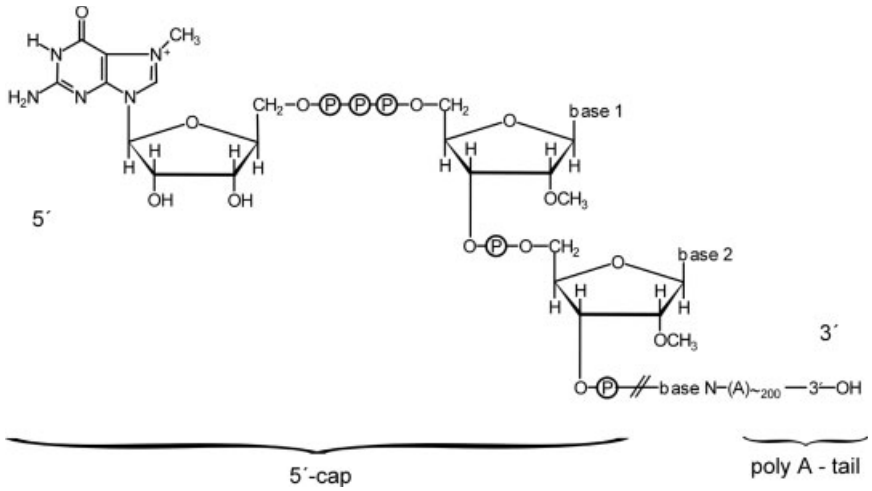


Fig. 1.40 Modifications at the 5' and 3' ends of eucaryotic mRNA. Eucaryotic mRNAs possess a "cap structure" at their 5' ends and a 100-200 base long poly-A tail at their 3' ends.

In higher eucaryotes, the 3' end of mature mRNA is not produced as a result of termination of transcription. Rather, the 3' end of the primary transcript is cut at a specific site and a poly-A sequence is appended. *Polyadenylation* precedes the splicing of the primary transcript.

During polyadenylation, the primary transcript is shortened in an endonucleolytic step and ca. 200 A-residues are appended. The endonucleolytic incision requires two signal sequences on the pre-mRNA. A highly conserved AAUAAA sequence 10 – 30 nucleotides upstream from the hydrolysis site serves as one signal. Another signal in the form of a less well-conserved GU- or U-rich element upstream of the hydrolysis site. Both together constitute the *polyadenylation signal*. Polyadenylation occurs in a multiprotein complex, components of which are associated with elongating RNA polymerase II.

1.5.2
Formation of Alternative mRNA by Alternative Polyadenylation and by Alternative Splicing

Some pre-mRNAs carry several polyadenylation signals on their 3'-end. Depending upon which signal is used for the polyadenylation, various mRNAs can be formed from a single primary transcript (Fig. 1.41). The mechanism of alternative polyadenylation offers the possibility to form cell- and tissue-specific mRNAs from the same primary transcript. It is still unkown how the cell decides which poly-A signal to use.

The major path for extracting different information from the same pre-mRNA uses alternative splicing. The genetic information encoding a protein in higher eucaryotes is usually found in pieces of coding sequences, or exons, interrupted by non-coding sequences, the introns. For the formation of the mature mRNA, the introns must be

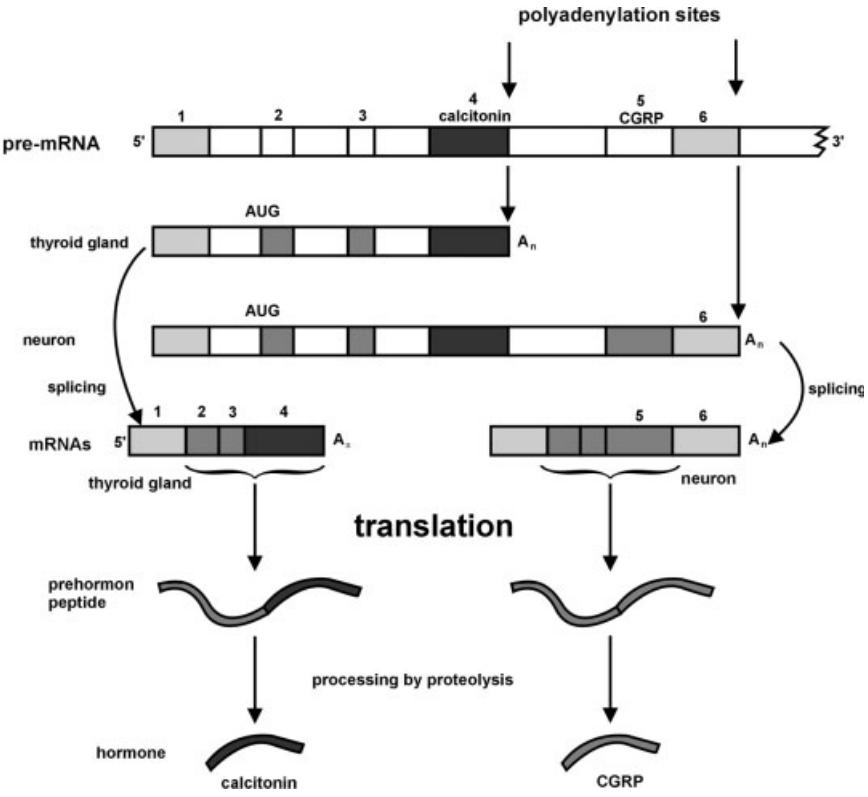


Fig. 1.41 Alternative polyadenylation in the expression of calcitonin genes of rat. The primary transcript of the calcitonin gene possesses two polyadenylation sites. One site is used in the processing of RNA in the thyroid, another site in the brain, and yet another in nerve tissue. The translation of the two mRNAs creates two pre-hormones, from which two different polypeptide hormones (calcitonin and the "calcitonin-related peptide", or CGRP) are created via proteolysis.

excised and the exons rejoined in the proper order. This process is termed splicing. The number of introns in eucaryotic genes can be very large; there are 50 introns in the human dystrophin gene.

Splicing occurs in a large protein-nucleic acid complex, termed the spliceosome. Components of the spliceosome are, apart from the pre-mRNA, a number of proteins and small RNAs, termed the U1, U2, U4, U5 and U6. The RNAs found in the spliceosome are bound to specific proteins. The complexes are termed snRNPs (small nuclear ribonucleoproteins). Depending upon the type of RNA bound, there are U1, U2, U5 and U4/U6 snRNPs.

Based on the observation of self-splicing by the 23S RNA of *Tetrahymena*, it is assumed that the cleavage and rejoining of the phosphodiester bond is catalyzed by the RNA components of the spliceosome. The proteins of the spliceosome are believed to be important for the recognition of the 5' and 3' splice sites and for the formation of a

defined structure in the spliceosome. Thus, the proteins of the spliceosome play a decisive role in the choice of the splice site and the efficiency of splicing.

The occurrence of several exons and introns in a gene opens up the possibility of *alternative splicing*.

Starting with a single pre-mRNA, several alternative mRNAs via the rejoining of various exons can be formed, each coding for proteins with different activities and functions. The creation of different proteins due to alternative splicing is a commonly used tool for enhancing the information contained within a gene and regulating its expression in a cell- and tissue-specific way. Some alternative splicing events appear to be constitutive, whereas others are regulated in response to developmental or physiological signals. In this way, a flexible adjustment of specific mRNA levels in response to a transduced external signal can be achieved. An example of the alternative splicing is shown in Fig. 1.42.

Alternative splicing can create proteins with varying enzymatic activity, cellular localization and regulatory activity. By including or excluding stop codons during alternative splicing, the expression of a protein can even be turned on and off. Misdirection of splicing has been recognized as a frequent cause of inheritable genetic diseases.

The splicing of pre-mRNAs often allows several possible combinations of exons. Despite this fact, only a few of the theoretically possible combinations are realized.

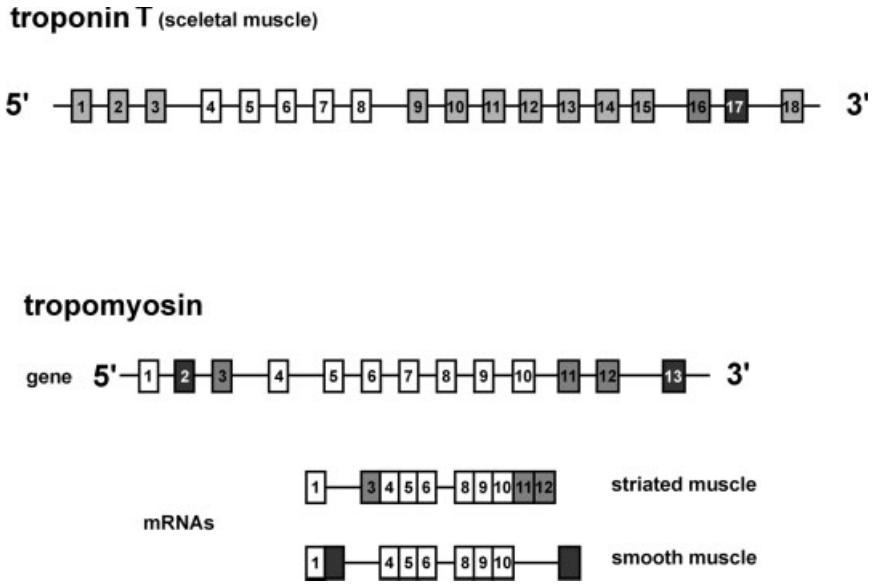


Fig. 1.42 Differntial splicing in muscle proteins. a) The troponin gene of rats possess 18 exons that encode 258 amino acids. Different subtypes of troponin are found in various types of muscle tissue.The exons 1-3 and 9-15 are found in all subtypes, while the exons 4-8 appear in various combinations, allowing 32 possible combinations. Exon 16 or 17 are found in every subtype. Altogether, 64 different mRNAs can be formed from the troponin pre-mRNA. b) Tropomyosin is a muscle protein that can be alternatively spliced in different muscle tissue. Shown in the figure are the predominant subtypes for striated and smooth muscle.

This accentuates the fact that the choice of alternative splice sites must be strictly controlled. The exact mechanism by which a splice site is chosen is not yet understood. The following mechanisms have been found to contribute to alternative splicing (review: Lopez, 1998):

1. ***Cis*-acting sequence elements**
   The choice of a particular splice site can be motivated by sequence elements that function as splicing enhancers. These enhancers, some of which are purine-rich, are assumed to bind and recruit the essential splicing factors.
2. **Proteins acting in *trans***
   Several proteins have been identified which act as antagonists in the selection of splice sites. The SF2 protein (and related proteins) belong to the family of SR proteins and support the use of 5' splice sites. SR proteins contain an RNA recognition motif (RRM) and are rich in the amino acids serine and arginine. Another protein, the hnRNP A1 protein supports the use of 3' splice sites.

Positive and negative regulatory proteins have been identified that help to enhance or suppress the use of a specific splice site by binding to cognate RNA elements. A positive-acting protein would accordingly stimulate the use of a specific splice site, while a negative-acting protein would inhibit the use of a splice site. Both processes allow for timely control of splicing and can effect a short-term change in gene expression in response to external signals. In addition, the activity of some of the proteins involved in splice site choice seems to be regulated by specific phosphorylation. Protein kinases which are specific for SR proteins and which can influence the subcellular localization and the splicing activity of these proteins by phosphorylating the SR motifs have been identified.

1.5.3
Regulation via Transport and Splicing of Pre-mRNA

Splicing and transport of the transcript from the nucleus to the cytosol are tightly coupled. Unspliced pre-mRNA usually does not leave the nucleus, so that only correctly spliced mRNA reaches the cytosol. The human immunodeficiency virus (HIV) is an example of how manipulation of the splicing pattern can be used to create different mRNAs from one pre-mRNA.

Following integration into the host genome as a provirus, the genome of a retrovirus is similar in many respects to a typical eucaryotic cellular gene. The transcription of the HIV coding DNA, modification and processing of the HIV transcript are performed by most of the same protein complexes which process cellular mRNAs. Among the proteins encoded by the HIV genome, some are required at the beginning of the replication cycle, while others are required at a much later phase. The regulatory proteins Tat and Rev are required at the beginning. Only when sufficient levels of HIV RNA have been produced are structural proteins, such as Gag, Pol and Env, required. The primary transcript of the HIV genome is a 9 kb long pre-mRNA, out of which further

RNAs arise by differential splicing (Fig. 1.43). The remarkable thing about the splicing of HIV RNA is the fact that the splicing and hence transport of the pre-mRNA is regulated by a viral protein.

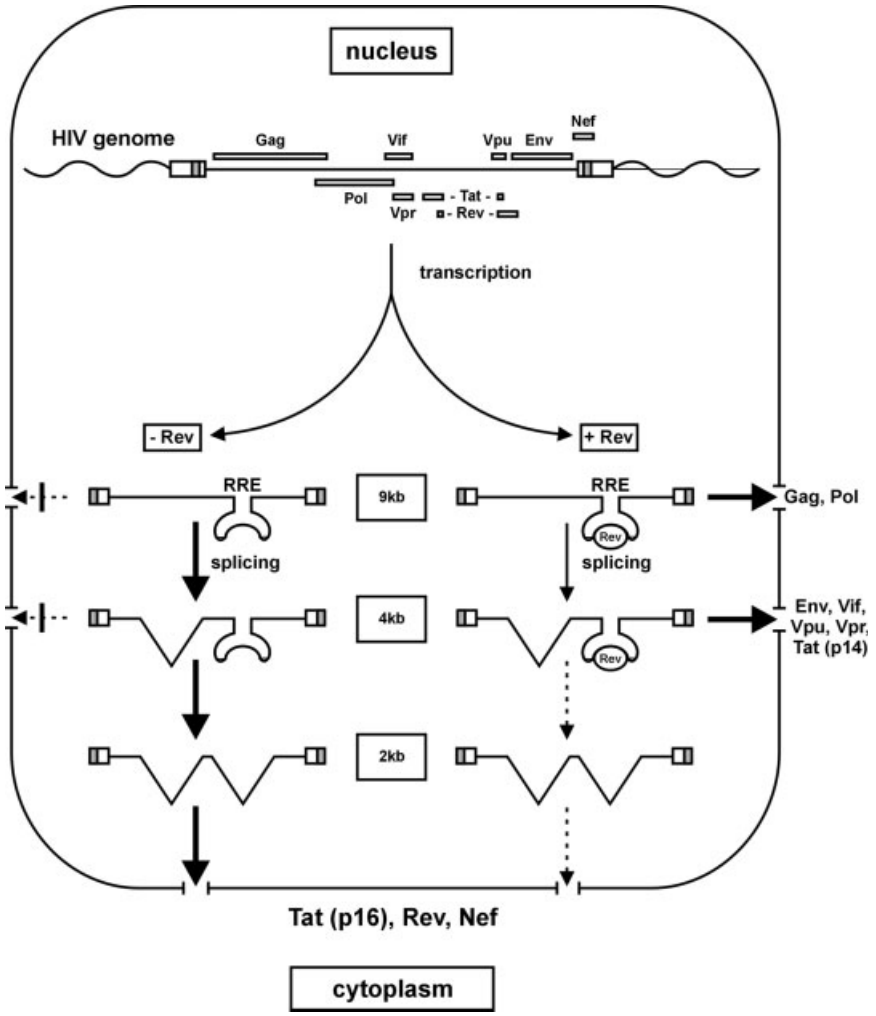


Fig. 1.43 The function of the Rev protein in the processing of the transcripts of the Human Immunodeficincy Virus (HIV). The Rev protein is a regulatory protein required for the processing of the primary transcripts of HIV. In the early stadium of viral replication, before Rev protein is available, spliced mRNA of ca. 2kb are transported to the cytosol. The spliced mRNAs are created by multiple splicing events of the primary transcript and encode the regulatory proteins Tat, Rev and Nef. Once a critical level of Rev protein is formed, unspliced (9kb) or only singly spliced (4kb) forms of the viral mRNA appear in the cytosol. These encode for structural proteins and reverse transcriptase. The Rev protein binds to a particular sequence of the viral RNA, the Rev responsive element (RRE). The binding of Rev to the RRE enables the transport of unspliced viral transcript into cytosol.

A key element of regulation is the Rev protein. In the absence of the Rev protein, one finds a multiply spliced class of 2S RNA in the cytoplasm that encodes the regulatory proteins Tat, Rev and Nef. At this phase, the mRNAs which encode structural protein are transported to the cytosol only at a low level. They are, furthermore, unstable and are only partially loaded with polysomes.

In the presence of the Rev protein, however, mainly longer, unspliced 9S mRNAs and partially spliced viral RNAs are found in the cytosol, from which the expression of the structural proteins occurs.

At the beginning of the replication cycle of HIV, before the Rev protein is present, regulatory proteins are primarily formed. If, in the framework of this expression pattern, enough Rev protein is present, then unspliced or incompletely spliced viral mRNAs appear in the cytosol, and structural proteins are formed.

The critical point for the Rev protein is the processing of the 9S and 4S RNA. In the presence of Rev, the splicing of long mRNAs to the short 2S RNA is repressed and the unspliced mRNA forms can be transported from the nucleus to the cytosol. The Rev protein binds to a ca. 230 long RNA element, termed the "Rev Responsive Element" (RRE), which includes sequences on the 9S and 4S RNA. The binding of Rev to the RRE obviously allows the transport of the unspliced or partially spliced RNA into the cytosol.

The exact mechanism by which Rev interferes with the transport and splicing process is largely unknown. One possible scenario is that Rev interacts with components of the spliceosome, which leads to the release of splice factors that allow cytosolic transport without actual splicing.

1.5.4
Stability of the mRNA

The concentration of mRNA available for translation is determined by the rates of nuclear RNA synthesis, processing and export, and the rate of cytoplasmic mRNA degradation. The specific degradation of mRNAs plays an important role in cell- and tissue-specific gene expression (review: Guhaniyogi and Brewer, 2001). The stability of various mRNAs can vary significantly, with half-lives ranging from 20 min to 24 h in the same cell. It is this wide range of mRNA decay rates that plays an important role in the regulation of gene expression. On one hand, unstable transcripts often encode proteins that are involved in short-term signaling like, e.g., oncoproteins and cytokines. These proteins are often induced upon external (stress, hormonal, nutritional) or internal (developmental) stimuli. On the other hand, stable transcripts often encode constitutively required proteins with housekeeping functions.

The stability of an mRNA is determined by *cis*-acting elements as well as by *trans*-acting protein factors. Destabilizing sequence elements were identified by site-specific changes in the mRNA sequences. They can be found at the 5' end, in the coding region, or at the 3' end in the form of, e.g., A+U rich elements. It can be generally noted that the degradation of mRNA by nucleases is coupled with translation. It is suspected that the proteins required for degradation are brought to the RNA by the translation apparatus. The nature of these proteins is not yet clear.

Several decay pathways exist for mRNAs. The major type of decay pathway is dependent on the deadenylation of mRNAS; another pathway is independent of it. Furthermore, a pathway has been identified that targets mRNAS with nonsense codons and degrades aberrant RNAs.

The following discussion gives examples of control elements important for the stability and degradation of mRNA.

Destabilizing Sequences at the 3' Nontranslated End

Tubulin is a well-studied example of the interference of gene products with the stability of mRNA. Processed tubulin binds as a dimer to the growing tubulin chain on the ribosome. The binding of the tubulin facilitates the attack by endonucleases on the ribosome-bound tubulin mRNA and thereby initiates the degradation of the mRNA (Fig. 1.44). The goal of this regulation process is to prevent the formation of excess tubulin. If tubulin is in excess, then the degradation of its own mRNA is induced and the synthesis of more tubulin is prevented.

Regulation of mRNA Stability by Iron

Two proteins are important for iron metabolism in mammalian cells: the *transferrin receptor* (TFR) and *ferritin*. Ferritin is a protein for the storage of iron. The production and its level are increased when more iron is available.

TFR is instrumental in the uptake of iron under conditions of low iron concentration. The concentration of cellular TFR is inversely correlated with the level of iron: if little iron is present, then the TFR concentration is increased, and if high levels of iron are present, the TFR concentration is decreased.

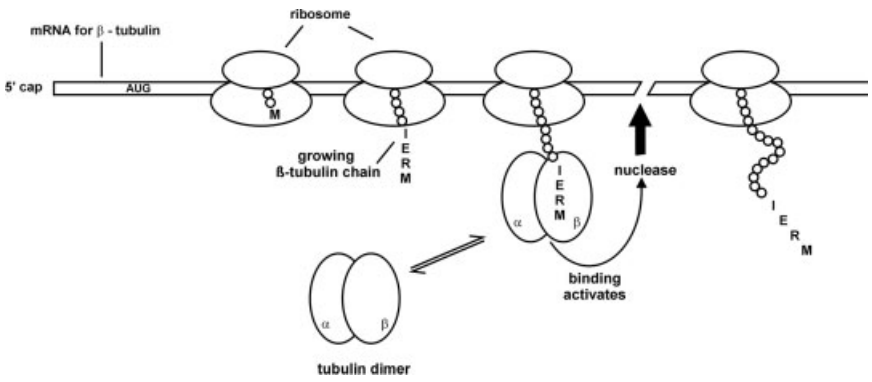


Fig. 1.44 Model for the control of translation by tubulin. The amount of tubulin in animal cells is determined partially by the stability of β-tubulin mRNA, whereby tubulin itself acts as the regulating signal. Starting from the 5' cap, various stages of the translation of β-tubulin mRNA, represented as a chain of small circles, is illustrated in the figure. As soon as the N-terminus of the growing β-chain emerges from the ribosome, the α- and β- subunits of tubulin bind to the terminal MREI sequence, upon which an endonuclease becomes activated by a presently unknown mechanism. The degradation of the β-tubulin mRNA then proceeds.

The regulation of the TFR and ferritin concentrations occurs at the mRNA level for both proteins (review: Theil and Eisenstein, 2000). The key element for the regulation of the TFR concentration is sequences at the 3' nontranslated end of the TFR mRNA. These sequences form hairpin structures that are essential for the iron-dependent stability of the mRNA. The hairpin structures are called iron-responsive elements (IREs) and offer binding sites for specific RNA-binding proteins, the iron regulatory proteins (IRP). Two types of IRPs are known, IRP1 and IRP2. Binding of IRPs to the hairpin structure protects the mRNA from nucleolytic degradation (Fig. 1.45).

The binding of the IRPs to the hairpin structures is controlled by the amount of iron in two ways. The IRP2 activity is regulated via an iron-induced protein oxidation, followed by ubiquitination and proteasomal degradation. By contrast, IRP1 is regulated via iron-dependent binding to the IREs. Low iron concentrations favor the formation of a binding-competent form of IRP1; high iron concentrations favor the formation of a binding-incompetent form. Both forms of the IRP1 differ from each other in terms of their content of a 4Fe-4S cluster. Iron favors the insertion of the 4Fe-4S cluster into the protein and thereby transmits it into the binding-incompetent state. In the presence of

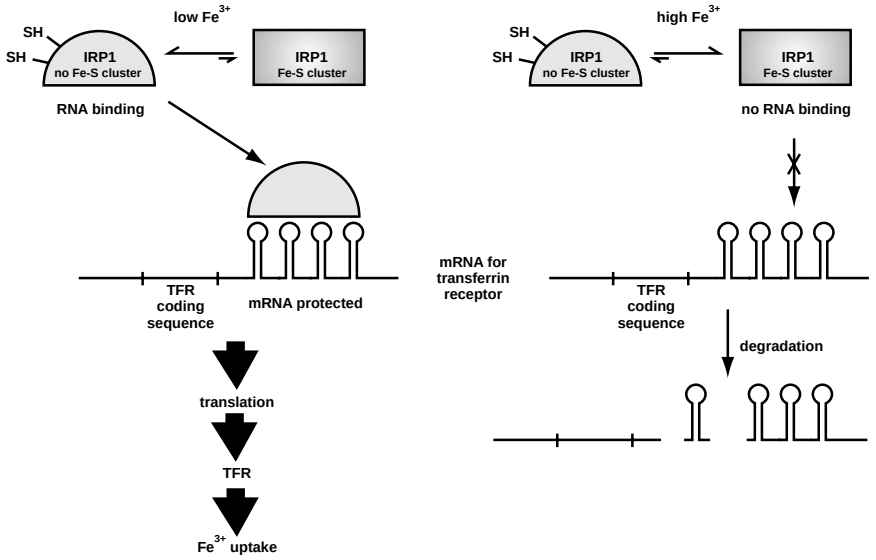


Fig. 1.45 Regulation of the stability of the mRNA of the transferrin receptor by Fe^{3+}. The translation of the mRNA of the transferrin receptor (TFR) is subject to regulation by the Fe concentration. Fe exerts its regulatory effect via the iron regulatory protein type1 (IRP1). The IRP1 binds to a control segment at the 3' terminal region of the TFR mRNA, known as the iron responsive element (IRE). Binding of IRP1 to a hairpin structure of the IRE stabilizes the mRNA of the transferrin receptor and protects it against nuclease attack. If high levels of Fe are present, then the IRP1 is in its binding inactive form, the IRE is free and nucleases can attack the non-translated 3' end of the mRNA. Upon depletion of Fe, the IRP1 is found in its binding compatible form and it protects the mRNA of the transferrin receptor from degradation. Binding-active and inactive forms of the IRP1 differ in the content of Fe-S clusters.

high levels of iron, the hairpin structures are not occupied, the mRNA can be degraded, and the level of TFR drops.

The regulation of the ferritin concentration is also related to the iron concentration. The vulnerable point is not the stability of the mRNA, but rather the initiation of translation. The mRNA for ferritin possesses an IRE element in the 5' non-coding region similar to that found in the TFR receptor at the 3' end. The IRP1 binds in its active form to the hairpin structure of the ferritin mRNA and blocks accessibility of the ribosome to the coding sequence. The translation of ferritin mRNA is halted under this mechanism upon low levels of iron, and the amount of ferritin decreases. At high iron concentrations the IRP1 is found in its binding-incompetent form, the 5' region is free to be translated, and *de novo* synthesis of ferritin for the storage of iron is possible.

1.5.5 Regulation at the Level of Translation

In order to allow a better understanding of regulation at the level of translation, some of the specific features of eucaryotic translation will be summarized briefly.

As opposed to the case in procaryotes, eucaryotic translation does not require a specific sequence for the binding of the ribosome. Procaryotes rely on the *Shine-Dalgarno sequence*, which is complementary to sequences of the 16S RNA of the 30S subunit. The Shine-Dalgarno sequence mediates the binding of mRNA to the 30S ribosome and ensures the correct positioning of the AUG initiation codon.

The mRNA of eucaryotes does not possess specific initiation sequences. Rather, the AUG start codon is identified by scanning the eucaryotic mRNA: the 40S subunit of the ribosome threads the 5' non-translated end of the mRNA and uses the first AUG codon encountered to initiate translation. Whether an AUG codon is used as an initiator depends, additionally, upon the sequence context. If the sequence environment is unfavorable for initiation, then the scanning is continued and initiation occurs at one of the next AUGs. With the help of this "leaky scanning" strategy, it is possible to produce proteins with different N-termini from the same mRNA. Since there are often signal sequences found at the N-terminus, this mechanism may lead to alternative compartmentalization of a protein.

Eucaryotic translation is controlled, analogously to transcription, primarily via initiation. Regulation by attenuation, an elementary regulation mechanism in procaryotes, is unknown to eucaryotes. Regulation by attenuation demands a tight coupling of transcription and translation. In eucaryotes such a coupling does not exist, since transcription and translation are spatially separated. As discussed above, however, eucaryotes possess an extensive repertoire of regulation possibilities in the framework of transport and processing of mRNA.

1.5.5.1 Regulation by binding of protein to the 5' end of the mRNA

As shown in the example of ferritin mRNA, the binding of the ribosome to the 5 end of the mRNA can be blocked by RNA-binding proteins at the same region. This case is an example of a negative regulation of translation by sequence-specific RNA-binding proteins. Such proteins are categorized as translation repressors (see Fig. 1.46).

Repression can be exerted by protein binding to the non-translated 5' end or the 3'-end of the mRNA. 5' end binding of e.g. IRP1 to the ferritin mRNA interferes with assembly of the initiation factors and the 40S subunit of the ribosome leading to translation inhibition. The mechanism of inhibition by by protein binding to the 3' end is not well understood. It is assumed that the 3' and 5' ends of the mRNA interact during the processes of capping, polyadenylation and 60S subunit binding. This interaction is

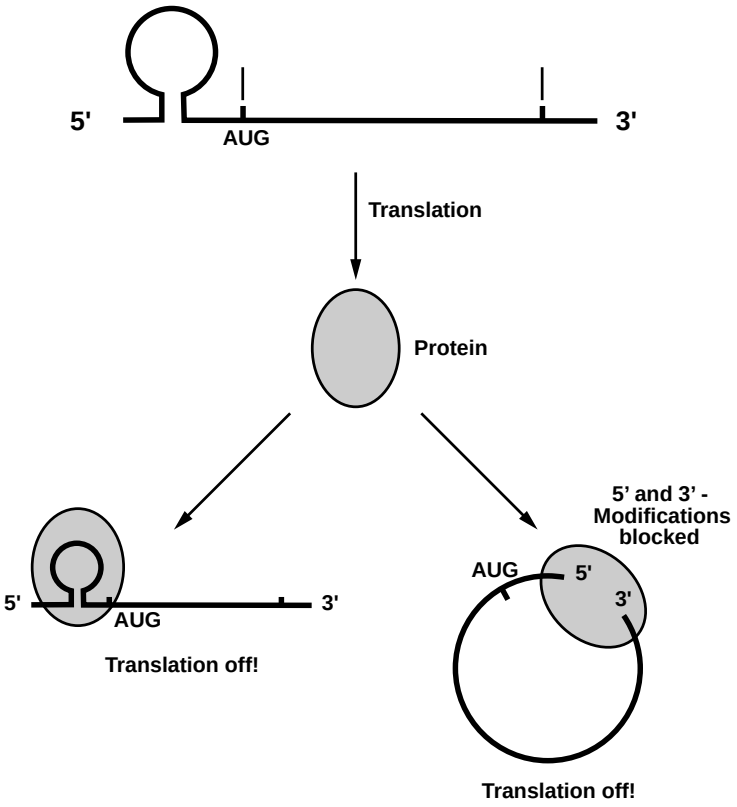


Fig. 1.46 Principle of negative control of translation initiation by protein binding to mRNA. Proteins can negatively effect translation by binding to the sequences in the 5' or 3' non-translated region of their own or other mRNAs. The participating proteins are sequence-specific RNA binding proteins and recognize RNA sequences in hairpin structures or other secondary structures of RNA at the 5'- and/or 3'-end. The protein binding to the 5'-end will interfere with the scanning of ribosomes whereas protein binding to the 3'-end is thought to prevent translation of mRNA via backfolding of the mRNA.

thought to be mediated by proteins bound to the 3' end of the mRNA and it is assumed that the binding of translation repressors to the 3' end interferes with these reactions.

Often repressor proteins are either directly or indirectly related to the protein encoded by the regulated mRNA, thus allowing adaptation of the translation process to the current needs by a feedback mechanism. The gene product then acts as a repressor of its own translation. This mechanism is utilized by procaryotes to regulate the synthesis of ribosomal proteins: excess ribosomal proteins block the translation of its own mRNA.

Translational control by 3' end binding is often used in developmental processes to inactivate a specific mRNA at certain developmental stages. Examples are the Pumilio protein (review: Parisi and Lin, 2000) and the fragile X mental retardation protein (review: Kaytor and Orr, 2001) which are sequence-specific RNA binding proteins that bind to short RNA motifs at the 3-end of specific mRNA thereby inhibiting mRNA translation.

An interesting case of regulation at the translation level was discovered in the case of the homeodomain protein *bicoid* (bcd), which is important in *Drosophila* differentiation (Dubnau and Struhl, 1996). The bicoid protein is a transcriptional activator that binds a cognate DNA element and stimulates the transcription of the neighboring genes. Apart from its specific DNA-binding capability, the bicoid protein binds to 3', non-translating sequences of the mRNA of another homeodomain protein (caudal protein) to inhibit its translation.

1.5.5.2 **Regulation by Modification of Initiation Factors**

The *de novo* synthesis of proteins can be varied over a wide range in response to external stimuli. Treatment of cells with hormones, mitogens or growth factors generally leads to an increase of protein biosynthesis. Conversely, lack of nutrients or environmental stresses like heat, UV irradiation or viral infections generally inhibit translation. The regulation of protein biosynthesis occurs primarily via phosphorylation of translation initiation factors. The regulatory points in eucaryotes are, above all, the translation factors eIF-2 and eIF-4 (reviews: Preiss and Hentze, 1999; Dever, 2002).

Regulation via eIF-2

The function of eIF-2 is illustrated schematically in Fig. 1.47. The translation factor eIF-2 belongs to the super family of regulatory GTPases (see Chapter 5), and fulfills the task of bringing the methionyl-initiator-tRNA to the 40S subunit of the ribosome. The active eIF-2*GTP form binds the methionyl-initiator-tRNA, associates with the cap structure of the mRNA, then commences to scan along the mRNA. Once an AUG codon is encountered, the bound GTP is hydrolyzed to GDP, resulting in the dissociation of the eIF-2*GDP from the 40S ribosome. The transition from the inactive eIF-2*GDP form into the active eIF-2*GTP form requires a G-nucleotide exchange factor, termed the eIF-2B which is composed of five polypeptides that can be divided into a regulatory subcomplex (subunits α,β,δ) and a catalytic subcomplex (ε,γ). The regulatory subcomplex interacts with eIF2α in a phosphorylation dependent manner.

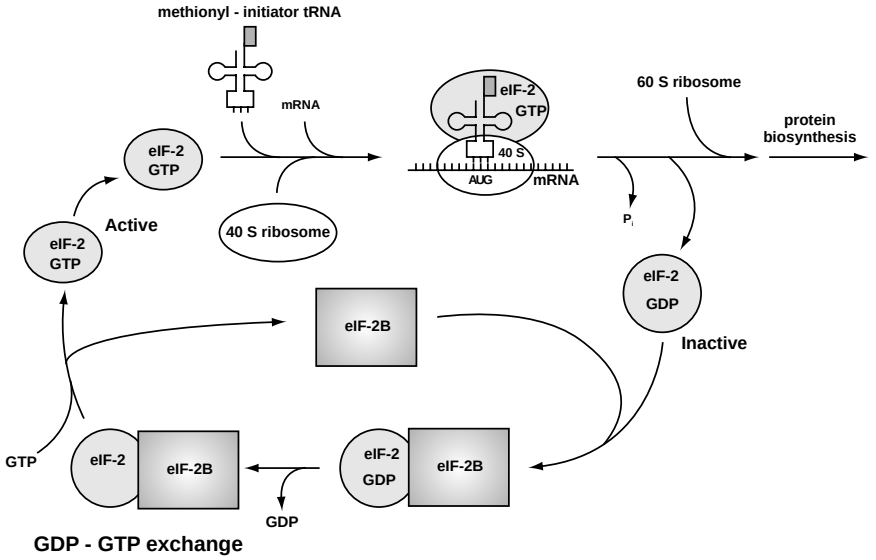


Fig. 1.47 The function of eIF-2 in eucaryotic translation. The initiator protein for translation, eIF-2, is a regulatory GTPase that occurs in an active GTP-form and in an inactive GDP form (see Chapter 5). The active eIF-2*GTP forms a complex with the initiator-tRNA, fMet-tRNA^{fmet} and the 40S subunit of the ribosome. This complex binds to the cap structure of mRNA to initiate the scanning process. eIF-2 undergoes an activation cycle typical for regulatory GTPases: the inactive eIF-2*GDP fom is activated with the assistance of the eIF-2B protein into the active eIF-2*GTP form. eIF-2B acts as a G-nucleotide exchange factor in the cycle (see Chapter 5).

In response to external signals, the α-subunit of eIF-2 is phosphorylated on Ser 51. This phosphorylation converts eIF2 from a substrate to an inhibitor of eIF2B. The affinity of eIF2 for the nucleotide exchange factor eIF-2B is increased, without inducing nucleotide exchange. Translation factor eIF-2 is found in excess of eIF-2B in the cell, so that phosphorylated eIF-2 binds the entire reservoir of eIF-2B. As a consequence, no further eIF-2B is available for nucleotide exchange, and protein biosynthesis is halted (see Fig. 1.48).

Four different protein kinases have been identified that specifically phosphorylate eIF2 on Ser 51 (review Dever, 2002). The eIF2α kinases comprise the families of HRI, PKR, PERK and GCN2 kinases.

The protein kinase HRI (heme regulated eIF-2 kinase) was first identified in studies on the regulation of protein biosynthesis in erythroid cells. A decrease in the heme concentration in reticulocytes leads to inhibition of globin synthesis at the level of translation. This regulation mechanism ensures that only so much globin is produced as is heme available. If the level of heme drops, then HRI becomes activated. The activated HRI phosphorylates the eIF-2α subunit, which in turn shuts off protein biosynthesis (Fig. 1.48). The mechanism of regulation of HRI kinase by heme is not well understood. Heme binding sites have been identified on the N-terminus and the kinase domain of HRI.

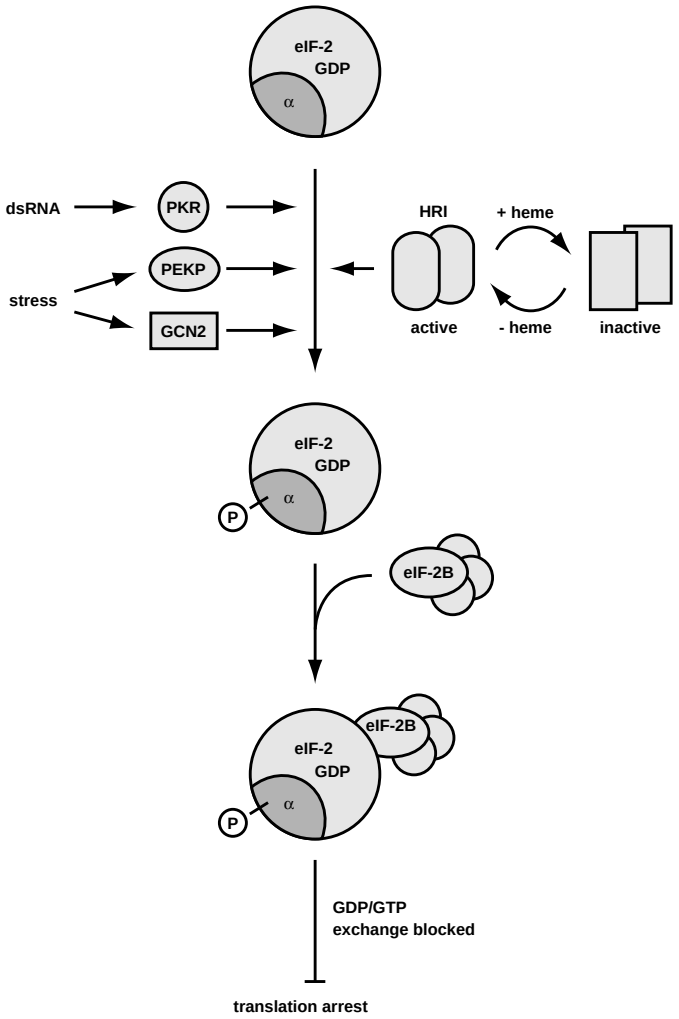


Fig. 1.48 Control of eIF-2 by phosphorylation. Phosphorylated eIF-2 · GDP binds strongly to the eIF-2B complex without necleotide exchange occurring. Initiation of protein biosynthesis is not possible in this case. At least four different protein kinases control the phosphorylation state of eIF-2 · GDP.
In reticulocytes, eIF-2 is subject to phosphorylation by the heme-regulated eIF-2-kinase (HRI) which is regulated via the heme concentration. Another protein kinase that can phosphorylate and regulate eIF-2 is the RNA-dependent eIF2α-kinase (PKR). The latter is induced by interferons and activated by double stranded RNA. Stress influences activate the protein kinases PRPK and GCN2 allowing for inhibition of protein synthesis via eIF-2 phosphorylation too.

The protein kinase PKR (RNA-specific eIF-2 kinase) is regulated by binding of dsRNA and by interferon on the level of expression. PKR contains two dsRNA binding sites and it is thought that dsRNA binding disrupts inhibitory interactions in PKR leading to its activation. The activation of PKR by dsRNA and its induction by interferon identify PKR as a component of the cellular anit-viral defense. Consistent with this notion, a large number of viruses express inhibitors of PKR.

The third eIF2α kinase, PERK participates in the endoplasmatic reticulum stress response. The fourth eIF2α kinase, GCN2, is activated under conditions of amino acid starvation.

Importantly, inhibition of translation via phosphorylation by the eIF2α kinases can have both a general and a gene-specific effect and can even lead to enhanced translation of specific mRNAs. Whereas the general level of translation may be reduced under these conditions, specific mRNAs are preferentially translated. This upregulation of specific mRNAs is explained by a leaky scanning of AUG codons and the use of alternative initiation sites.

Regulation via eIF-4: Influence of Insulin and Nutrients

It has long been known that insulin (and other hormones and growth factors) can stimulate protein biosynthesis. On the other hand, deprivation of nutrients and stress exposure can reduce protein biosynthesis. The signal transduction pathway linking external signals like insulin to the translation apparatus was for a long time unclear. However, insight is now being gained into this mechanism (review: Dever, 2002).

The critical target for insulin-mediated regulation of translation is the initiation factor eIF-4E. This factor binds specifically to the 5'-cap structure of mRNA as part of a larger complex, termed eIF-4F, composed of eIF-4E, eIF-4A, which possesses RNA helicase activity, and eIF4G. Binding of eIF-4E to the eIF-4F complex is necessary for eIF-4F recruitment to the cap structure and for the subsequent steps of translation initiation, namely for the association of the 40S subunit with the 5'-end of the mRNA and for the scanning by the 40S subunit. The availability of a functional eIF-4F cap-binding complex is controlled by a specific binding protein, 4E-BP, which competes with eIF-4G for a common binding site on eIF-4E. The binding of 4E-BP to eIF-4E is regulated by phosphorylation of 4E-BP. In the hypophosphorylated form, 4E-BP binds strongly to eIF-4E and thereby outcompetes eIF4G. In this situation, a functional eIF-4F complex cannot form, the 5' cap structure is not bound and translation of mRNA is inhibited. Upon signal-mediated phosphorylation of a critical set of 5-6 Ser/Thr residues on 4E-BP, its interaction with eIF-4E is abrogated, eIF-4G binds to eIF4-E, the eIF-4F complex forms at the cap structure and the translation block is relieved (Fig. 1.49).

The phosphorylation of 4E-BP1 is under the control of at least two signaling pathways, the Akt-pathway and the mTOR-pathway, as discussed below.

Under the influence of insulin (and other hormones or growth factors), a signaling chain is activated that results in the activation of the serine/threonine-specific protein kinase Akt/PKB (see Section 6.6.3) and in the phosphorylation of 4E-BP1. It is assumed that the phosphorylation of 4E-BP1 represents the signal for the release of

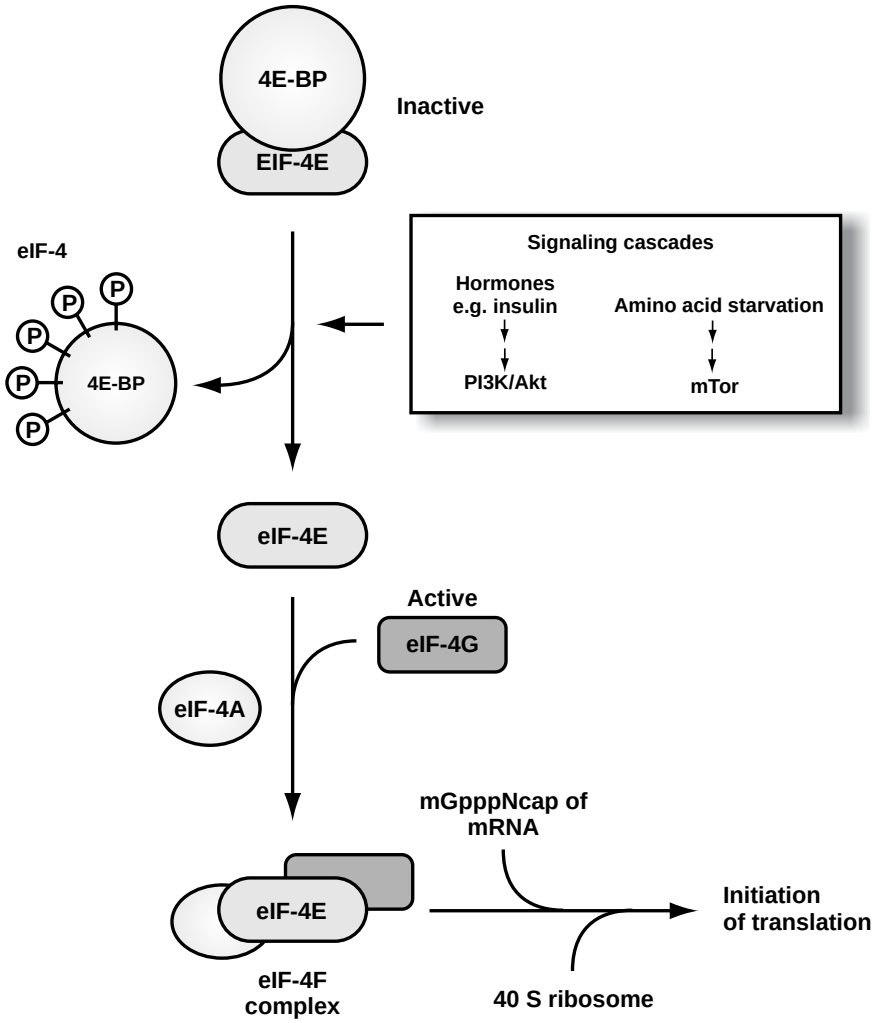


Fig. 1.49 Model of the regulation of translation by insulin. Insulin (and other growth factors) activates the Akt kinase pathway (see Chapter 6), whose final result is the phosphorylation of 4E-BP1, a regulatory protein of translation initiation. The multiply phosphorylated 4E-BP1 protein strongly binds to the initation factor eIF-4E making it unavailable for translation initiation. eIF-4E is required, together with the proteins eIF-4A and eIF-4G, for the binding of the 40S subunit of the ribosome to the cap structure of the mRNA. If the 4E-BP1 protein becomes phosphorylated as a result of insulin-mediated activation of the PI3-kinase/Akt kinase cascade, then eIF-4E is liberated from the eIF-4E · 4E-BP1 complex, the ternary complex between eIF-4E, eIF-4A and eIF-4G forms and protein biosynthesis can begin.

eIF-4E from the binary complex. The free eIF4E can then associate with eIF-4A, eIF4G, and the 5' cap of the mRNA to initiate translation (Fig. 1.49).

Nutrients influence translation via another pathway, which also has eIF-4E as a target. When sufficient nutrients are available, the serine/threonine-specific protein kinase mTOR (mammalian target of rapamycin; review: Gingras et al., 2001) is activated. mTOR belongs to the class of PI3-like protein kinases and was named according to its specific binding to the complex between the immunosuppressant rapamyacin and the FK504-binding protein (see also Section 7.6.5). Active mTOR generates a signal that results in the phosphorylation of 4E-BP1 and translation stimulation. Lack of amino acids like leucine leads to a reduction of mTOR activity and of 4E-BP1 phosphorylation and hence inhibition of translation.

The regulation of translation is accomplished in the two systems via a specific inhibitory protein and an initiation factor of translation. The binding activity of the inhibitor protein is regulated by protein phosphorylation, and thus by protein kinases. *Signals from diverse signaling pathways may use different protein kinases to achieve phosphorylation of the same target protein.*

A further susceptible point for both the insulin-Akt signaling pathway and the mTOR pathway is the ribosomal protein S6. Under the influence of insulin, S6 is phosphorylated by a specific protein kinase, the $p70^{S6}$ kinase (review: Avruch et al., 2001), resulting in increased levels of translation of certain mRNAs. Several pathways, including the MAPK/ERK pathway (see Chapter 10) and the Akt kinase pathway, can contribute to the activation of the $p70^{S6}$ kinase.

Reference

1 Agalioti, T., Lomvardas, S., Parekh, B., Yie, J., Maniatis, T. and Thanos, D. (2000) Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell*, **103**, 667–678.

2 Albright, S.R. and Tjian, R. (2000) TAFs revisited: more data reveal new twists and confirm old ideas. *Gene*, **242**, 1–13.

3 Avruch, J., Belham, C., Weng, Q., Hara, K. and Yonezawa, K. (2001) The p70 S6 kinase integrates nutrient and growth signals to control translational capacity. *Prog. Mol. Subcell. Biol.*, **26**, 115–154.

4 Ballestar, E. and Wolffe, A.P. (2001) Methyl-CpG-binding proteins. Targeting specific gene repression. *Eur. J. Biochem.*, **268**, 1–6.

5 Berger, S.L. (2001) Molecular biology. The histone modification circus. *Science*, **292**, 64–65.

6 Boyes, J., Byfield, P., Nakatani, Y. and Ogryzko, V. (1998) Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature*, **396**, 594–598.

7 Carrion, A.M., Link, W.A., Ledo, F., Mellstrom, B. and Naranjo, J.R. (1999) DREAM is a Ca2+-regulated transcriptional repressor. *Nature*, **398**, 80–84.

8 Chan, H.M. and La Thangue, N.B. (2001) p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J. Cell Sci.*, **114**, 2363–2373.

9 Chang, W.H. and Kornberg, R.D. (2000) Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH. *Cell*, **102**, 609–613.

10 Chen, F.E., Huang, D.B., Chen, Y.Q. and Gosh, G. (1988) Chrystal structure of p50/p65 heterodimer of transcription factor NF-κB bound to DNA. *Nature*, **391**, 410–413,

11 Cramer, P., Bushnell, D.A. and Kornberg, R.D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*, **292**, 1863–1876.

12 Dever, T.E. (2002) Gene-specific regulation by general translation factors. *Cell*, **108**, 545–556.

13 Driever, W. and Nusslein-Volhard, C. (1989) The bicoid protein is a positive regulator of hunchback transcription in the early Drosophila embryo. *Nature*, **337**, 138–143.

14 Dubnau, J. and Struhl, G. (1996) RNA recognition and translational regulation by a homeodomain protein. *Nature*, **379**, 694–699.

15 Egly, J.M. (2001) The 14th Datta Lecture. TFIIH: from transcription to clinic. *FEBS Lett.*, **498**, 124–128.

16 Ghosh, G., van Duyne, G., Ghosh, S. and Sigler, P.B. (1995) Structure of NF-kappa B p50 homodimer bound to a kappa B site. *Nature*, **373**, 303–310.

17 Gingras, A.C., Raught, B. and Sonenberg, N. (2001) Regulation of translation initiation by FRAP/mTOR. *Genes Dev.*, **15**, 807–826.

18 Goll, M.G. and Bestor, T.H. (2002) Histone modification and replacement in chromatin activation. *Genes Dev.*, **16**, 1739–1742.

19 Guhaniyogi, J. and Brewer, G. (2001) Regulation of mRNA stability in mammalian cells. *Gene*, **265**, 11–23.

20 Hovde, S., Abate-Shen, C. and Geiger, J.H. (2001) Crystal structure of the Msx-1 homeodomain/DNA complex. *Biochemistry*, **40**, 12013–12021.

21 Karimpour, S., Lou, J., Lin, L.L., Rene, L.M., Lagunas, L., Ma, X., Karra, S., Bradbury, C.M., Markovina, S., Goswami, P.C. *et al.* (2002) Thioredoxin reductase regulates AP-1 activity as well as thioredoxin nuclear localization via active cysteines in response to ionizing radiation. *Oncogene*, **21**, 6317–6327.

22 Karin, M., Liu, Z. and Zandi, E. (1997) AP-1 function and regulation. *Curr.Opin.Cell Biol.*, **9**, 240–246.

23 Kaytor, M.D. and Orr, H.T. (2001) RNA targets of the fragile X protein. *Cell*, **107**, 555–557.

24 Kim, Y., Geiger, J.H., Hahn, S. and Sigler, P.B. (1993) 'Crystal structure of a yeast TBP/DNA complex' *Nature 365*, 512–520

25 Kobor, M.S. and Greenblatt, J. (2002) Regulation of transcription elongation by phosphorylation. *Biochim. Biophys. Acta*, **1577**, 261–275.

26 Kouzarides, T. (2002) Histone methylation in transcriptional control. *Curr. Opin. Genet. Dev.*, **12**, 198–209.

27 Lewis, M., Chang, G., Horton, N.C., Kercher, M.A., Pace, H.C., Schumacher, M.A., Brennan, R.G. and Lu, P. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science*, **271**, 1247–1254.

28 Lomvardas, S. and Thanos, D. (2001) Nucleosome sliding via TBP DNA binding in vivo. *Cell*, **106**, 685–696.

29 Lopez, A.J. (1998) Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu. Rev. Genet.*, **32**, 279–305.

30 Luisi, B.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R. and Sigler, P.B. (1991) Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature*, **352**, 497–505.

31 Malik, S. and Roeder, R.G. (2000) Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem.Sci.*, **25**, 277–283.

32 Marmorstein, R. and Roth, S.Y. (2001) Histone acetyltransferases: function, structure, and catalysis. *Curr. Opin. Genet. Dev.*, **11**, 155–161.

33 Merika, M. and Thanos, D. (2001) Enhanceosomes. *Curr. Opin. Genet. Dev.*, **11**, 205–208.

34 Mowen, K.A., Tang, J., Zhu, W., Schurter, B.T., Shuai, K., Herschman, H.R. and David, M. (2001) Arginine methylation of STAT1 modulates IFNalpha/beta-induced transcription. *Cell*, **104**, 731–741.

35 Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O. and Darst, S.A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*, **296**, 1285–1290.

36 Myers, L.C. and Kornberg, R.D. (2000) Mediator of transcriptional regulation. *Annu. Rev. Biochem.*, **69**, 729–749.

37 Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D. and Grewal, S.I. (2001) Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science*, **292**, 110–113.

38 Nakayama, T. and Takami, Y. (2001) Participation of histones and histone-modifying enzymes in cell functions through alterations in chromatin structure. *J. Biochem.(Tokyo)*, **129**, 491–499.

39 Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. and Bird, A. (1998) Transcriptional repression by

the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, **393**, 386–389.

40 Narlikar, G.J., Fan, H.Y. and Kingston, R.E. (2002) Cooperation between complexes that regulate chromatin structure and transcription. *Cell*, **108**, 475–487.

41 Parisi, M. and Lin, H. (2000) Translational repression: a duet of Nanos and Pumilio. *Curr.Biol.*, **10**, R81–R83.

42 Pavletich, N.P. and Pabo, C.O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science*, **252**, 809–817.

43 Proudfoot, N.J., Furger, A. and Dye, M.J. (2002) Integrating mRNA processing with transcription. *Cell*, **108** , 501–512.

44 Ptashne, M. and Gann, A. (1997) Transcriptional activation by recruitment. *Nature*, **386**, 569–577.

45 Richards, E.J. and Elgin, S.C. (2002) Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell*, **108**, 489–500.

46 Romano, G., Kasten, M., De Falco, G., Micheli, P., Khalili, K. and Giordano, A. (1999) Regulatory functions of Cdk9 and of cyclin T1 in HIV tat transactivation pathway gene expression. *J. Cell Biochem.*, **75**, 357–368.

47 Shaulian, E. and Karin, M. (2002) AP-1 as a regulator of cell life and death. *Nat.Cell Biol.*, **4**, E131–E136.

48 Tan, S. and Richmond, T.J. (1998) Eukaryotic transcription factors. *Curr. Opin. Struct. Biol.*, **8**, 41–48.

49 Thomas, J.O. and Travers, A.A. (2001) HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem. Sci.*, **26**, 167–174.

50 Tong, J.K. (2002) Dissecting histone deacetylase function. *Chem. Biol.*, **9**, 668–670.

51 Veenstra, G.J. and Wolffe, A.P. (2001) Gene-selective developmental roles of general transcription factors. *Trends Biochem. Sci.*, **26**, 665–671.

52 Werner, M.H. and Burley, S.K. (1997) Architectural transcription factors: proteins that remodel DNA. *Cell*, **88**, 733–736.

2
The Regulation of Enzyme Activity

Central to cell regulation is the ability of a cell to adopt the activity of enzymes in response to external and internal signals. All functions of a cell are based on the fine-tuned action of enzymes that are part of regulatory networks receiving and producing regulatory signals. The rates of the biochemical reactions within an organism are precisely adjusted to nutrient supply and to the requirements of growth and division of each cell, thereby ensuring survival of the whole organism. How much of a specific substrate within a cell is converted to the product is determined by the *actual enzyme activity* that catalyzes this reaction. How active a given enzyme is at a given time and at a specific location within the cell is subjected to a multitude of regulatory mechanisms that operate at different levels and with different goals.

In a broad sense, enzyme regulation operates at two main levels, either by changing the concentration of an enzyme or by altering the activity of *preexisting* enzymes.

The *concentration* of enzymes can be regulated, e.g., via
- transcription, translation and processing of the mRNA (see Chapter 1)
- targeted proteolysis
- targeted translocation.

The activity of preexisting enzymes can be regulated via
- binding of effector molecules with a positive or negative influence on enzyme activity
- covalent modification e.g. phosphorylation
  acetylation
  methylation

Cells use regulation at the level of enzyme concentrations often for long-term adaptations during cell growth and cell differentiation where a fast biochemical response to a regulatory signal is not required. Furthermore, by fine-tuning the rates of degradation and synthesis of an enzyme a distinct steady state concentration of an enzyme can be established.

In contrast, regulation via modulation of the activity of preexisting enzymes is often used for a quick response to internal or external stimuli, allowing short-term changes in the rates of biochemical reactions.

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

2.1 Enzymes as Catalysts

Enzymes function as biocatalysts and, as such, are involved in all metabolic reactions. Characteristics of enzymes are their high *efficiency*, high *specificity*, extreme *stereoselectivity*, and their *ability to be regulated*. Analogously to chemical catalysts, enzymes do not alter the equilibrium of a reaction, but only accelerate the establishment of that equilibrium.

The mechanism of the action of enzymes can best be described with the aid of transition state theory. On the pathway from substrate A to product B in a reaction catalyzed by a chemical catalyst or an enzyme, A passes through a transition state $A^{\#}$, which is found at the highest point of the energy diagram (Fig. 2.1). The energy difference between the ground state of A and the transition state $A^{\#}$ represents the activation energy. The transition state as such cannot be isolated. It is the state of A in which the bonds participating in the reaction are in the process of opening and closing. The transition state is the most unstable state on the path from substrate to product. Enzymes, like chemical catalysts, increase the rate of a reaction by decreasing the activation energy for the transition from A to B.

For the tight binding of the transition state, the binding surface of the enzyme must be complementary to the structure of the transition state, so that optimal interactions between the enzyme and the transition state are possible. This demand implies that

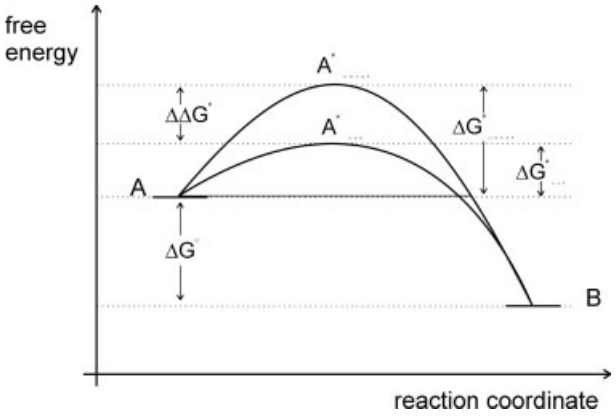


Fig. 2.1 The energy profile of a catalyzed and an uncatalyzed reaction. The figure shows the energy diagram for the conversion of A->B for a catalyzed and an uncatalyzed reaction. The binding of A to the catalyst (=enzyme) is left out to simplify the figure. In the uncatalyzed reaction, the energy difference between the ground state A and the transition state $A^{\#}$ is the activation energy $G^{\#}_{uncat}$. The transition state of an catalyzed reaction is at a lower free energy, so that the activation energy $G^{\#}_{cat}$ is less and the reaction proceeds at a faster rate. The energy difference $G^{\#}$ is a measure for how much faster the catalyzed reaction is compared to the uncatalyzed reaction. The equilibrium of the reaction, which is characterized by ΔG_0, remains unchanged upon catalysis.

enzymes display a high affinity to molecules which are chemically similar to the transition state of the reaction. Complexes of such transition state analogs with enzymes are well suited for X-ray structure analysis to elucidate the structural principles of the active site and the catalytic mechanism.

The pathway from enzyme-bound substrate to the transition state involves changes in the electronic configuration and geometry of the substrate. The enzyme itself is also not static. The ability to tightly bind the transition state requires flexibility in the active site. Such flexibility has been experimentally demonstrated in many cases. A corollary to this is that the effectivity of enzyme catalysis can easily be influenced and regulated by conformational changes in the enzyme. An extensive consideration of the mechanisms of enzymes can be found in the works by J. Kraut (1988) and A. Fersht (1998).

The binding of an effector molecule or a covalent modification of the enzyme, such as phosphorylation, can prevent the restructuring of the enzyme that is necessary for strong binding of the transition state, and the reaction is inhibited. Effectors and enzyme modification can also affect the substrate-binding site such that binding of the substrate in the ground state is impossible or very weak. On the other hand, activation by effectors and modifications can be achieved by stabilizing a conformation of the enzyme in which substrate binding is favorable and a high complementarity between enzyme and substrate in the transition state is possible.

The ability of proteins to exist in different conformations is termed allostery (see Section 2.3). Allosteric enzymes can assume various conformations, which differ in catalytic activity and/or substrate-binding capacity.

2.2
Regulation of Enzymes by Effector Molecules

The regulation of the activity of enzymes by the binding of effector molecules is a ubiquitous and general principle for the fine tuning and control of metabolic activity. Effector molecules are often low-molecular-weight organic compounds. Proteins and metal ions can also exercise the function of effectors. The effector molecules bind specifically to the enzymes, and the binding results in inhibition or stimulation of enzymatic activity.

For the regulation of metabolic pathways, metabolites are often used which are a product of that pathway. The basic strategy for the regulation is exemplified in the mechanisms employed in the biosynthetic and degradation pathways of amino acids, purines, and pyrimidines, as well as in glycolysis. In most cases a metabolite (or similar molecule) of the pathway is utilized as the effector for the activation or inhibition of enzymes in that pathway.

Effector molecules bind to an enzyme of a metabolic pathway and modify its activity in a concentration-dependent manner. This regulation serves to adjust the level of production of metabolites by the pathway to the current needs in the cell. The regulatory signal of such a "*feedback regulation*" strategy is the concentration of the final product or of an intermediate. In feedback inhibition, if the concentration of a particular product exceeds a certain level, then the product occupies the active site of the

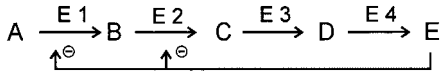


Fig. 2.2 Feedback inhibition of a metabolic pathway. In feedback inhibition, the final product, E, of a metabolic path acts as an inhibitor of an early reaction in a cycle, which in the figure leads eventually to the formation of E with the aid of enzymes E1, E2, etc. E acts as an effector on the enzymes E1 and E2 to inhibit their activity.

enzyme and thereby inhibits it. The enzymes of a metabolic pathway that act as a bottleneck for the total throughput, i.e. are rate limiting for the overall reaction, are generally the primary targets for feedback regulation. In this case the allosteric regulation of one enzyme can control the overall throughput of the metabolic pathway. These enzymes are recognized as *key enzymes* and are usually involved in early steps in the reaction path (Fig. 2.2).

The basic elements for the regulation of enzyme activity by effector molecules are *allosteric conformational changes* of the enzyme. Allostery means that the enzyme can exist in various conformations which differ in activity and substrate or ligand binding. The typical titration curve for the binding of a ligand to an allosteric protein is sigmoidal in shape (Fig. 2.3), while binding curves for nonallosteric enzymes are hyperbolic.

Enzymes that are regulated by effector molecules in an allosteric manner possess, apart from the binding site for the substrate, a specific binding site for the effector molecule. The binding of effector molecules to the effector site leads to a shift in equilibrium between the various conformations and thus to a change in activity (see Section 2.3).

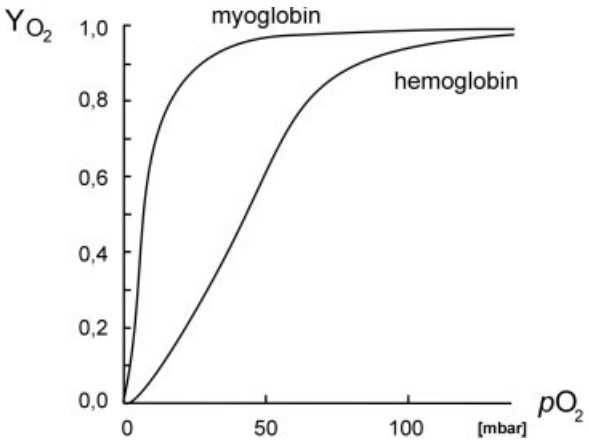


Fig. 2.3 Ligand binding to allosteric and non-allosteric proteins: oxygen bindig to myoglobin and hemoglobin. a) the binding of O_2 to myoglobin can be described by a hyperbolic curve. b) the sigmoid form of the O_2 binding curve to hemoglobin is characteristic for a ligand binding to an allosteric protein.Y= degree of binding, ratio of binding sites occupied with O_2 to the total O_2 binding sites of hemoglobin; pO_2: partial pressure of O_2.

2.3
Principal Features of Allosteric Regulation

Mechanistic studies and model considerations have revealed the following general characteristics of allosterically regulated enzymes (for details see Fersht, 1998):

1. Allosteric proteins are typically composed of two or more, often symmetric subunits.
2. The subunits can exist in active and inactive forms, often termed the T-form and the R-form. The R-form ("relaxed") is the relaxed, active state; the T-form ("tense") is the less active state. T and R forms are in equilibrium with each other.
3. The ligands can bind to both the T and R forms. The two forms differ in their affinity to the ligand. Ligand binding can often be described by a hyperbolic binding curve (see Fig. 2.3).
4. Effector molecules that function as *activators* bind preferentially to the R form of the enzyme and thereby stabilize it. *Inhibitors* bind preferentially to the T form. In the presence of an inhibitor, more molecules occur in the T form. In this form the enzyme possesses a lower affinity for the substrate and the enzyme is thus less active.

Fig. 2.4 illustrates the influence of an activator and of an inhibitor on the binding of a ligand (and thereby on the activity) by a tetrameric protein. The activator shifts the

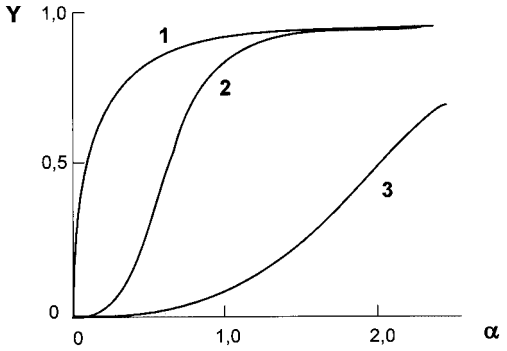


Fig. 2.4 The influence of an activator and an inhibitor on the ligand binding curve of an allosterically regulated enzyme. In the absence of inhibitor or activator, the ligand binding curve is sigmoidal (curve 2). In the presence of an activator the binding curve is shifted to lower ligand concentrations (curve 1) and approaches hyperbolic form, similar to that observed in Fig. 2.3 for the binding of O2 to myoglobin. An inhibitor shifts the binding curve to higher ligand concentrations (curve 3). Y: degree of binding, ratio of occupied to toal binding sites of the protein; a: L_0/P_0; L_0: total ligand concentration; P_0: total concentration of binding sites on the protein.

binding curve to lower ligand concentrations and can, in the extreme case, lead to a hyperbolic binding curve. The inhibitor shifts the binding curve to higher ligand concentrations, so that higher concentrations are required to saturate the ligand binding sites.

The molecular basis for the allosteric regulation of enzyme activity has been studied in detail for many systems (see textbooks). Some general conclusions can be drawn from these studies:

- The interface between the subunits often plays a central role in allosteric regulation by allowing the coupling of conformational changes within one subunit to the other.
- Binding sites for substrates are frequently found at or near the interface between the subunits. The mutual orientation of the subunits thereby can influence the accessibility of the substrate binding sites.
- Active and inactive forms differ in the accessibility of the substrate binding site and/or the proper positioning of the catalytic residues.
- The extent of the conformational changes involved can be variable. Long-range, medium-range and short-range conformational changes have been described for allosteric enzyme regulation.
- The binding of an effector molecule to the effector binding site induces conformational changes that influence – often via the subunit interfaces – the substrate binding site and/or the orientation of the catalytic residues. Thereby, the enzyme is fixed in either the active or the inactive state.

2.4 Regulation of Enzyme Activity by Binding of Inhibitor and Activator Proteins

Enzyme-specific inhibitor and activator proteins can be considered as a type of effector molecules.

Inhibitor Proteins

There are numerous examples of inhibitor proteins that specifically bind a particular enzyme and block its activity.

Inhibitors can use a variety of mechanisms to control enzyme activity:

- **Binding to the substrate binding site**
  Inhibitors may be related structurally to the substrate without possessing the chemical groups that are necessary for being turned over. Because of the structural similarity of the substrate, these inhibitors can bind specifically to the substrate binding site and compete with the substrate for the enzyme.
  The structurally best characterized example of this type of inhibition is found amongst the proteases: specific inhibitor proteins ensure that the proteases carry out their proteolytic function only in certain tissues and only under certain metabolic conditions. By this token, the protease inhibitors assume an important pro-

tective and regulatory function in metabolism. The structural analysis of the trypsin inhibitor from bovine pancreas (BPTI) in complex with trypsin shows that the inhibitor occupies and blocks the substrate-binding pocket in a highly complementary manner (see textbokks for details). The inhibitor can be likened to a *pseudosubstrate* and, as such, is bound with high affinity. The cleavage of the peptide bond is, however, not possible because the structural requirements for enzymatic cleavage are not fulfilled.

- **Deformation of the active site and/or the substrate binding site**
  Binding of the inhibitor to the enzyme may alter the orientation of the catalytically essential groups in a way that does not allow efficient catalysis and/or strong binding of the substrate. Examples are the inhibitors of the cyclin-dependent protein kinases of the cell cycle (see Chapter 13).

Inhibitor proteins themselves are subject to a variety of regulation mechanisms. The function of an inhibitor protein can be regulated, for example, by protein phosphorylation (see Section 7.6.2:), by degradation or by *de novo* synthesis (see Chapter 13).

Activator Proteins

Examples of the reversible association of activator proteins with an enzyme are the Ca^{2+}-calmodulin- dependent enzymes. Calmodulin is a Ca^{2+}-binding protein which can activate target enzymes, e.g., phosphorylase kinase (see Section 6.7.1 and Section 7.5) in its Ca^{2+}-bound form. Other examples of activating proteins are the cyclins (see Chapter 13). The cyclins are activators of protein kinases that regulate the cell cycle. Structural studies have shown that the cyclins activate the protein kinases of the cell cycle by restructuring the active site and positioning the catalytic groups to allow efficient catalysis.

Activator proteins themselves can be bound in regulatory networks, as shown in the example of the cyclins (Chapter 13). The function of an activator protein can be regulated at the level of gene expression, degradation, or post-translational modification (e.g., protein phosphorylation).

Metal Ions

The availability of metal ions can also be employed for regulation of enzyme activity. Of primary importance is Ca^{2+}. An important example in this regard is protein kinase C, which is activated by Ca^{2+} (see Section 7.4). The availability of Ca^{2+} is further regulated in various ways by hormone-controlled pathways (see Chapter 6).

2.5 Regulation of Enzyme Activity by Phosphorylation

The phosphorylation of enzymes by specific protein kinases is a widespread mechanism for the regulation of enzyme activity. It represents a flexible and reversible means of regulation and plays a central role in signal transduction chains in eucaryotes.

Proteins are phosphorylated mainly on *Ser/Thr* residues and on *Tyr* residues. Occasionally Asp or His residues are phosphorylated, the latter especially in procaryotic

signal transduction pathways (see Chapters 7 and 12). For the regulation of enzyme activity the phosphorylation of Ser and Thr residues is most significant. Apart from regulation of Tyr kinases, Tyr phosphorylation serves the function of creating specific attachment sites for proteins. Both of these functions will be discussed in more detail in Chapter 8.

Protein phosphorylation is a specific enzymatic reaction in which one protein serves as a substrate for a protein kinase. Protein kinases are phosphotransferases. They catalyze the transfer of a phosphate group from ATP to an acceptor amino acid in the substrate protein (Fig. 2.5). A detailed discussion of protein kinases can be found in Chapter 7.

The response of a protein upon phosphorylation is dictated by the special properties of the phosphate group. The phosphate group has a pK_a of about 6.7 and carries two negative charges at neutral pH. Therefore, two negative charges are generated in a substrate protein upon phosphorylation of an uncharged amino acid side chain. This fact and the presence of four oxygen atoms allows the phosphate group to form an extensive network of H-bonds which can link different parts of a polypeptide chain. Similar twofold negatively charged groups do not occur in other structural elements of proteins. Electrostatic interactions and a network of H-bonds are therefore of special importance for the control of protein functions by phosphorylation.

Analysis of protein-protein interactions in existing structures has shown that phosphate groups most commonly interact with the main-chain nitrogens at the start of a helix, where often glycine is found. In nonhelix interactions, phosphate groups most commonly contact arginine residues. The guanidinium group of arginine is well suited for interactions with phosphate because of its planar structure and its ability to form multiple hydrogen bonds. The electrostatic interaction between arginine residues and the phosphate group provides tight binding sites that often function as organizers of short-range as well as long-range conformational changes.

The phosphate esters of Ser, Thr, or Tyr residues are quite stable at room temperature and neutral pH; the rate of spontaneous hydrolysis is very low. Therefore, to remove the phosphate residue the cell utilizes specific enzymes termed phosphatases. Based on substrate specificity, these can be classified as Ser/Thr or Tyr-specific phosphatases (see Chapters 7 and 8).

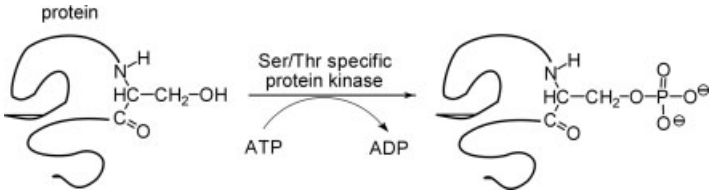


Fig. 2.5 Change in charge state of proteins via phosphorylation. The phosphorylation of Ser residues is catalyzed by a Ser/Thr-specific protein kinase that utilizes ATP as the phosphate group donor. The product of the reaction is a Ser-phosphate ester which carries a net charge of -2 at physiological pH.

Structural Consequences of Protein Phosphorylation

The molecular basis of the control function of protein phosphorylation will be illustrated for two enzymes, the glycogen phosphorylase from rabbit muscle and the isocitrate dehydrogenase from *E. coli*. Both examples represent very different mechanisms by which the activity of an enzyme is altered by protein phosphorylation. Glycogen phosphorylase displays extensive allosteric conformational changes upon protein phosphorylation. Isocitrate dehydrogenase is phosphorylated directly in the substrate-binding site with only minimal conformational changes resulting. For further examples on the structural aspects of the control by phosphorylation see Johnson and O'Reilly (1997).

2.5.1 Regulation of Glycogen Phosphorylase by Phosphorylation

Glycogen phosphorylase was the first enzyme shown to be regulated via protein phosphorylation (Krebs, 1959). In recognition of their trail-blazing work, Edwin Krebs and Edmond Fisher were awarded the Nobel prize for Chemistry in 1992.

Glycogen phosphorylase catalyzes the transfer of a glucose unit from glycogen to an inorganic phosphate to form Glc-1-P. This first step of glycolysis is the target of several regulatory processes. The phosphorylase is subject to allosteric regulation by effector

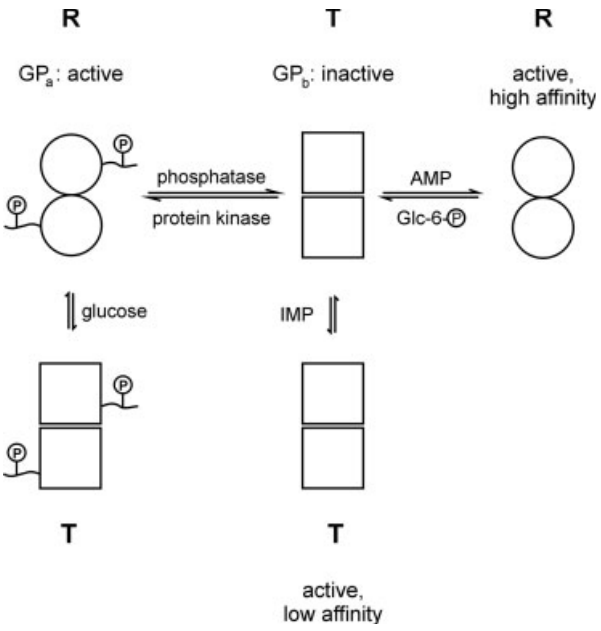


Fig. 2.6 Model of allosteric and covalent activation of glycogen phosphorylase of muscle. The R-form of the subunits are represented by circles, the T-form by squares. The active state of glycogen phosphorylase (GP) is characterized by a high affinity, the inactive state by low affinity for the substrate P_i.

molecules as well as a reversible regulation via phosphorylation. The regulation of the phophorylase by phosphorylation is the last step in a hormone-controlled signal chain, the individual steps of which will be discussed in Chapters 5, 6 and 7.

In the framework of models for allosteric control(see Section 2.3), a T- and R-form can be formulated for glycogen phosphorylase. In the T-form, glycogen phosphorylase binds its substrates and activating effectors with lower affinity, while in the R-form it possesses higher affinity for substrates and activating effectors.

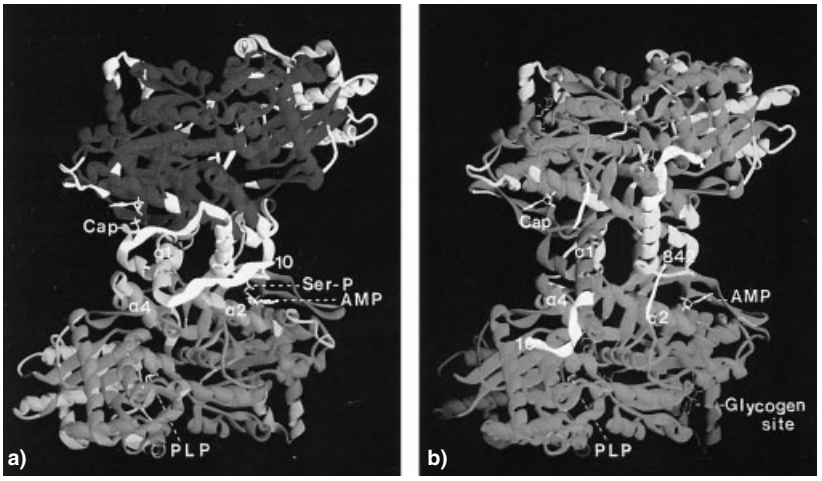
A simplified representation of the various conformational forms of glycogen phosphorylase is given in Fig. 2.6. The nonphosphorylated form is termed *phosphorylase b*. In the absence of effectors, phosphorylase b is found in the inactive T-form. Phosphorylation on Ser14 by phosphorylase kinase transforms phophorylase b into *phosphorylase a*, for which equilibrium lies toward the active, or R-form. The transition of phosphorylase b from the T-form to the R-form can also be induced by the binding of the activator AMP. Binding of the inhibitor Glc-6-P transforms the R-form of phosphorylase b back into the T-form. IMP and glucose have also been described as inhibitors that fix the enzyme in the T-state. In the absence of effectors, glycogen phosphorylase b occurs as a dimer. Binding of the activator AMP or phosphorylation leads to a partial aggregation of the dimers to tetramers. For structural discussions, however, it is sufficient to consider only the dimeric form of the enzyme.

High-resolution X-ray structures of the a- and b-form of rabbit muscle phosphorylase permit a view into some of the structural differences of the various allosteric forms of the enzyme. Furthermore, the data give an impression of the mechanism of binding of effectors and the influence that phosphorylation has on substrate binding and enzyme activity. The following discussion will be restricted to the observed consequences of phosphorylation.

Phosphorylation of the glycogen phosphorylase of muscle occurs on Ser14 near the N-terminus. In the unphosphorylated T-state, the amino acids 10–18 are found in disordered structures or irregular β-sheets and interact with other residues in the same subunit. In the T-form, the side chain of Ser14 is located near Glu501 of the same subunit. Phosphorylation of Ser14 leads to a radical restructuring (Fig. 2.7A and B): the N-terminus reorients itself in the opposite direction and interacts with the other subunit. The reorientation brings the negatively charged phosphate moiety closer to positively charged residues, substantially stabilizing the alternative conformation. Coupled to the restructuring of the N-terminus is a reorientation of additional structural elements, which eventually affects the binding site for the phosphate substrate. The final result is an active site conformation more favorable for substrate binding.

Phosphorylation of glycogen phosphorylase is the initiator for the coupled conformational changes, which are communicated over a large distance to the active site. Similar to the allosteric mechanism of phophofructokinase, the inter-subunit contact surfaces play a decisive role for the communication between the phosphorylation site and the substrate-binding site. They act as hinges for the transmission of the signal, allow interactions over large distances and provide for a cooperative coupling of conformation changes between the subunits. For a more detailed account one should refer to the original literature (Barford et al., 1991).

A



B

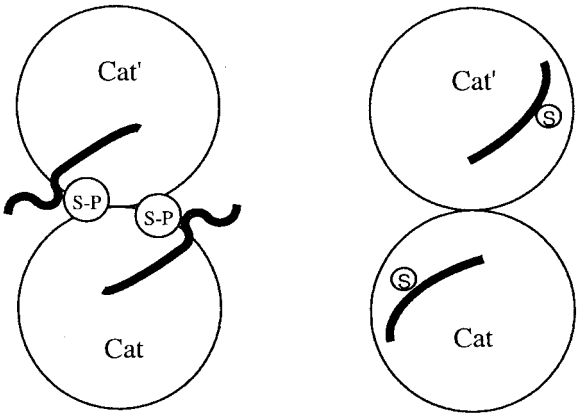


Fig. 2.7 A) Structural changes at the N-terminus of rabitt muscle glycogen phosphorylase as a result of phosphorylation. The N-terminal residues (10-23) and the C-terminal residues (837-842) are shown in white. a) R-form of the dimer of glycogen phosphorylase a. b) T form of the dimer of glycogen phosphorylase b. Phosphorylation at Ser13 near the N-terminus transforms the inactive glycogen phosphorylase b into the active glycogen phosphorylase a. The N-terminus rearranges significantly as a result of phosphorylation. In the inactive T-state, the N-terminus interacts with the same subunit, while in the R-form it forms interactions with the other subunit. After Barford and Johnson (1991), with permission.
B) Simplified schematic drawing showing changes in the N-terminus of rabitt muscle glycogen phosphorylase.The N-terminal tail is shown as a thick black line. The catalytic site Cat is on the far side of the molecule to the viewer.

2.5.2
Regulation of Isocitrate Dehydrogenase (*E. coli*) by Phosphorylation

Isocitrate dehydrogenase catalyzes the NAD-dependent reduction of isocitrate to α-ketoglutarate. The dimeric enzyme is regulated via phosphorylation. Phosphorylation on Ser113 leads to a complete inactivation of the enzyme.

An understanding of the molecular basis for regulation of isocitrate dehydrogenase by phosphorylation was facilitated by X-ray crystallography of the phosphorylated enzyme in complex with isocitrate. The crystal structures of mutants of the enzyme in which Ser113 had been exchanged for aspartate or glutamate were also solved (Hurley et al., 1990). The structure of the enzyme in complex with the substrate isocitrate revealed the phophorylation site to be localized near isocitrate. Ser113 itself binds the substrate directly via an H-bond with the O^4 of isocitrate (Fig. 2.8). The comparison of various forms of the enzyme demonstrated that phosphorylation does not always imply extensive conformational changes. The observed small conformational differences between the phosphorylated and unphosphorylated form may be irrelevant for the regulation mechanism.

Decisive for the inactivation of the enzyme via phosphorylation of Ser113 is the creation of negative charges at the substrate-binding site, which aggravates binding of the negatively charged isocitrate.

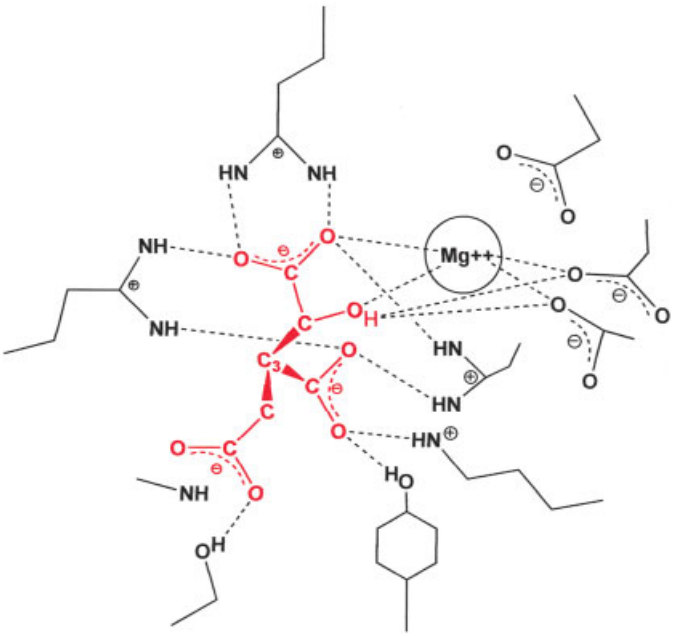


Fig. 2.8 Substrate binding site of isocitrate dehydrogenase from *E.coli*. The interactions involved in the binding of Mg^{2+}-isocitrate (red) at the active site of isocitrate dehydrogenase are shown. After Hurley et al., (1990).

As a consequence of electrostatic repulsion, the isocitrate substrate can only bind with low affinity to the phosphorylated enzyme. This interpretation is supported by structural data on mutant forms of isocitrate dehydrogenase. If one replaces Ser113 with amino acids that possess negatively charged side chains, then isocitrate can still bind, albeit with markedly reduced affinity, and the binding can only occur with an unfavorable geometry. Theoretical calculations of the change in free energy for the binding of isocitrate upon phosphorylation corroborates the conclusion that introduction of a negative charge to Ser113 is sufficient to explain the experimentally observed decrease in affinity for isocitrate.

2.6 Regulation via the "Ubiquitin-Proteasome" Pathway

Proteolysis is a versatile tool in the cell for the targeted change in structure, activity, function and subcellular distribution of proteins.

Examples of directed proteolysis are:

- removal of the N-terminal methionine residue from *de novo* synthesized proteins
- removal of the signal peptide upon protein translocation across the membrane
- hydrolysis of the translation product of monocistronic mRNA in viruses
- maturation of proteins: transformation of inactive proteins into active proteins
- targeted degradation of proteins.

Only the targeted degradation of proteins will be discussed in the following. For the maturation of proteins and for proteolysis by classical proteases the reader is referred to biochemical textbooks.

The function of most proteins in the cell is tightly restricted both locally and temporally. To what extent an enzyme can participate in metabolism depends not only on the rate of biosynthesis and on the extent of covalent modification, such as phosphorylation, but also on the rate of degradation of a protein. The comparison of the *life span* of various proteins (Table 2.1) shows that the stability of proteins can differ significantly. Among the most unstable, short-lived proteins are found many regulatory proteins and enzymes with key positions in a metabolic pathway. The targeted degradation of proteins involved in the cell cycle (see Chapter 13) is an essential element of cell cycle regulation.

Attempts to explain the variable life span of proteins showed that, apart from the nonspecific degradation of proteins, there also exist specific degradation mechanisms. Such specific mechanisms allow the function of a protein to be temporally restricted and specifically modified. There are two main pathways for the degradation of proteins in mammalian cells. In the lysosomal path, proteins that enter the cell via endocytosis are degraded. The degradation of proteins in the lysosome is rather unspecific and used mainly to eliminate foreign proteins. More recent findings indicate that a specific degradation pathway also exists in the lysosome (see below).

The nonlysosomal degradation pathways allow for the selective degradation of proteins under normal cellular conditions. These degradation pathways are also respon-

Tab. 2.1 Life span of selected proteins.

Life span (h)	Nucleus	Cytosol	ER and plasma membrane
< 2	Transcription factors Fos, Myc, Myb, p53, Hsp70	Hsp 70, protein kinase C Tyrosine aminotransferase	HMGCoA reductase
2–8	Cyclins	Protein kinase A Tryptophan oxygenase	γ-Glutamyl transferase
9–40	Ubiquitin#	Calmodulin, glucokinase, Ubiquitin	EGF receptor, epoxide hydrolase
41–200	Histone H1, HMG1, HMG2	Alcohol dehydrogenase, catalase Myoglobin, Phospholipase A2	Cytochrome b5, Dipeptidyl-dipeptidase
> 200	Histone H2A, H2B, H3, H4	Phosphoglycerate kinase Glycogen phosphorylase	NAD glycohydrolase, Acetylcholine receptor

sible for the degradation of cellular proteins under conditions of stress. The most significant and well-characterized nonlysosomal degradation pathway is that of the *ubiquitin-proteasome pathway* in which proteins are degraded in a 26S proteasome after they have been conjugated by one or more ubiquitin molecules. The ubiquitin-proteasome system (review: Hershko and Ciechanover, 1998; Pickart, 2001) is a tool for the selective proteolysis of proteins and thus plays an important regulatory role in the cell.

2.6.1 Components of the Ubiquitin System

Ubiquitin is a 76-residue protein found in nearly every eucaryote. It occurs either in free form or bound to other proteins. All known functions of ubiquitin are transmitted via its covalent linkage with other proteins. This serves the purpose, among others, of marking the proteins for proteolytic degradation.

The ubiquitinylation of proteins is a complex process, which involves three sequential enzymatic reactions performed by three types of enzymes, E1, E2 and E3 (Fig. 2.9). The enzymatic conjugating cascade is organized in a hierarchical way: There is one E1 enzyme, a limited number of E2 enzymes, each of which may serve several E3 enzymes, and a much larger number of E3 enzymes (Fig. 2.10).

Activation of Ubiquitin, Formation of E1-Ub

In an initial reaction, ubiquitin is activated by linking its C-terminal glycine carboxylate to an SH-group of the ubiquitin-activating enzyme E1. This step requires ATP and consists of an intermediate formation of ubiquitin adenylate followed by the binding of G76 of ubiquitin to a Cys residue of E1 in a thiolester linkage, with the release of PP_i and AMP.

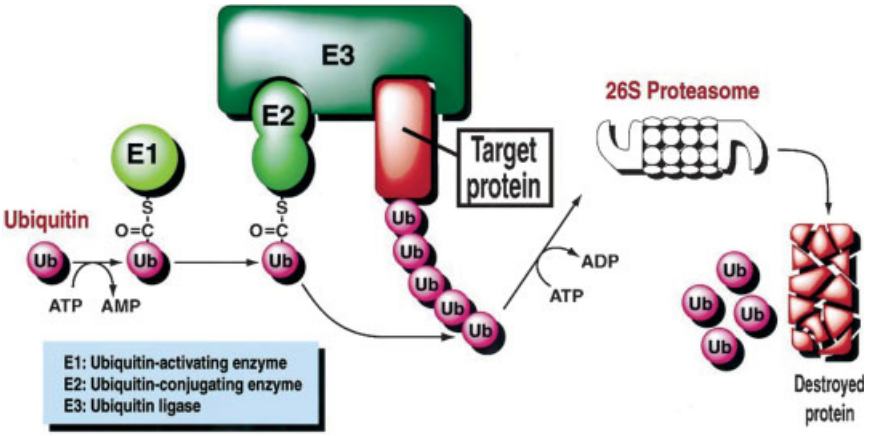


Fig. 2.9 Ubiquitinylation of proteins and degradation in the proteosome. Ubiquitin (Ub) is initially activated by an enzyme E1, whereby the C-terminal carboxyl group of ubiquitin becomes attached to an SH group of E1 via a thioester bond. The activated ubiquitin is then transferred from E1-Ub to the ubiquitin-conjugating enzyme, E2. Finally, the ubiquitin is covalently attached to the target protein in a reaction catalyzed by the E3 ubiquitin ligase. Repetition of this series of events results in the formation of a polyubiquitin chain on the target protein. The E3 enzyme binds directly to the target protein and therefore plays an important role in determining the substrate specificity of this pathway. The polyubiquitinylated protein is then degraded to peptides in the 26S proteosome. After Nakayama et al., (2001).

Transacylation to the Ubiquitin-conjugating Enzyme E2

In a transacylation reaction, the ubiquitin moiety is transferred from E1-Ub to a cysteine-SH within the active site of of the ubiquitin-conjugating enzyme E2 to form E2-Ub.

Transfer to the Target Protein with the Participation of E3

The third step of ubiquitinylation, the transfer of ubiquitin to the target protein, is catalyzed by a ubiquitin-protein-ligase or E3 enzyme. In this reaction ubiquitin is linked by its C-terminal glycine in an amide isopeptide linkage to an ε-NH_2-group of the substrate protein's Lys residues. In a subsequent reaction, the ubiquitin attached to the target protein can itself be a substrate for E3-mediated ubiquitinylation, resulting in the attachment of multiple ubiquitin molecules in a ubiquitin-ubiquitin linkage and the formation of poly-ubiquitin chains on the target protein. In these chains the ubiquitin molecules are linked by an isopeptide bond between K48 (or K29, K63) and G76. The presence of poly-ubiquitin chains on the target protein is a prerequisite for its degradation in the proteasome.

The E3 enzymes are primarily responsible for conferring specificity to ubiquitin conjugation. Each E3 enzyme, together with its cognate E2 enzyme, recognizes a few substrates that share a particular ubiquitination signal, which is usually a primary sequence motif. Subsequently the substrate is marked with a secondary signal, the poly-ubiquitin chain. The proteasome recognizes only this secondary signal and therefore will degrade a huge variety of substrates. Selection of substrates for ubiquitin

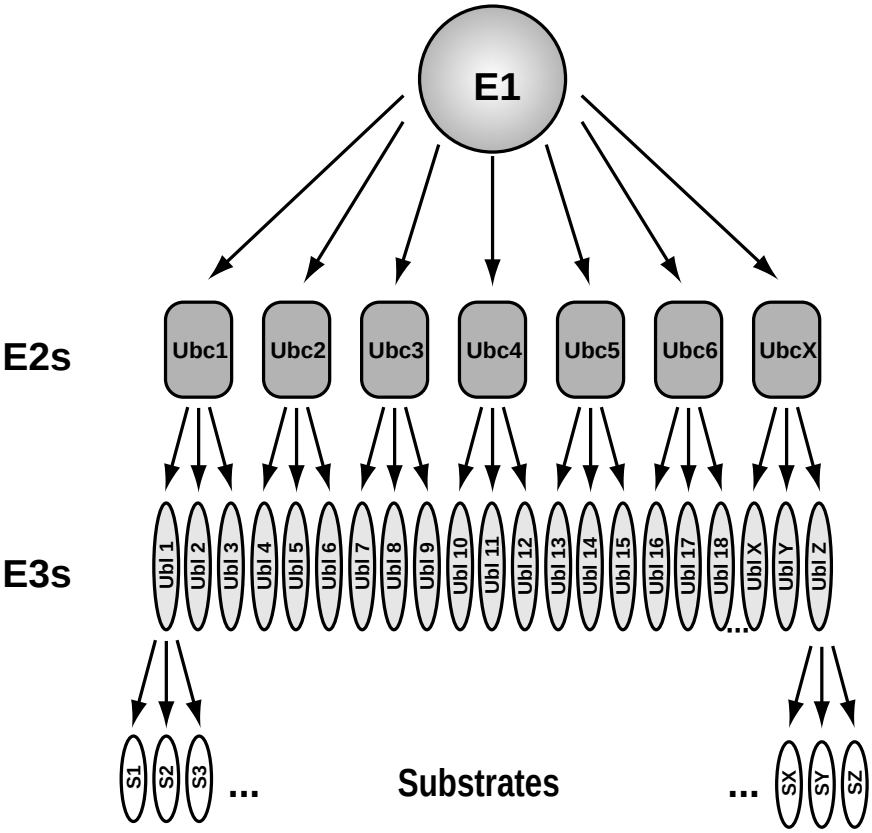


Fig. 2.10 Hierarchy of enzymes of ubiquitin conjugation. There is one E1 enzyme, several E2 enzymes and many E3 enzymes. Each of the many E3 enzymes has distinct set of proteins as substrates.

ligation and thus for proteasome degradation occurrs primarily by the E3 enzymes, of which a large number are known.

Ubiquitin can be transferred to the substrate protein by two different mechanisms:

- In some families of E3 enzymes, ubiquitin is first transferred from the E2 carrier to an active site cysteine of the E3 enzyme and subsequently to the ε-NH_2-group of an acceptor lysine on the substrate protein.
- In other families of E3 enzymes, no intermediate E3-ubiquitin linkage can be demonstrated. In this case, ubiquitin is transferred directly from E2 to the substrate protein. The E3 enzymes are nevertheless required for ubiquitinylation since the E3 enzymes are responsible for substrate selection and are found in tight complexes with the cognate E2 proteins.

Based on common structural features, E3 enzymes are classified into two families, the Hect domain E3 enzymes and the RING finger E3 enzymes.

- **Hect domain E3 enzymes**
  The discovery of this family of E3 enzymes started from the studies on the targeted degradation of the p53 tumor suppressor protein. Ubiquitinylation and degradation of p53 can be mediated by the papilloma virus E6 oncoprotein (see below) in collaboration with a further protein, E6-AP (E6 associated protein). E6-AP was the first member of a large family of E3 enzymes, the Hect (homologous to E6-AP C-terminus) domain family. These proteins form a ubiquitin-E3 intermediate and then transfer the ubiquitin to lysine side chains of the substrate protein. They contain an essential active site Cys residue near the C-terminus and one or several WW domains (see Chapter 8.2.6). Fig. 2.11 illustrates schematically the function of Hect – type E3 ligases on the example of the p53 ubiqutiniylation.
- **RING finger E3 enzymes**
  We now know of many E3 enzymes that share a series of histidine and cysteine residues with a characteristic spacing that allows for the coordination of two zinc ions in a structure called the Really Interesting New Gene (RING) finger. The RING finger of E3 enzymes is found in single-subunit enzymes like the Cbl-protein (see below) or in the form of multi-subunit E3 enzymes (SCF, APC, see below) where the RING finger is located in one specific subunit (reviews: Joazeiro and Weissman, 2000; Pickart, 2001). It is now generally assumed that RING finger motifs mediate the interaction between E3 enzymes and ubiquitin-conjugating E2 enzymes, thereby promoting ubiquitination of target proteins that have been specifically recognized by E3 enzymes. The RING finger does not directly participate in the ubiquitin transfer to the target protein, but rather seems to function as a scaffold that positions the substrate and the E2 enzyme optimally for ubiquitin transfer. As indicated schematically in Fig. 2.12 the RING finger E3s have to fulfill at least three tasks: the recognition and binding of the substrate, the recognition and binding of the E2 enzyme, and the participation – either directly or indirectly – in the transfer of the ubiquitin. For some types of E3 enzymes, substrate recognition is mediated primarily by a specific subunit containing a sequence motif called F-Box (see also Fig. 2.13) It is still an open question how the E3s act in the transfer step. Up to now, catalytic residues in the E3s have not been identified.

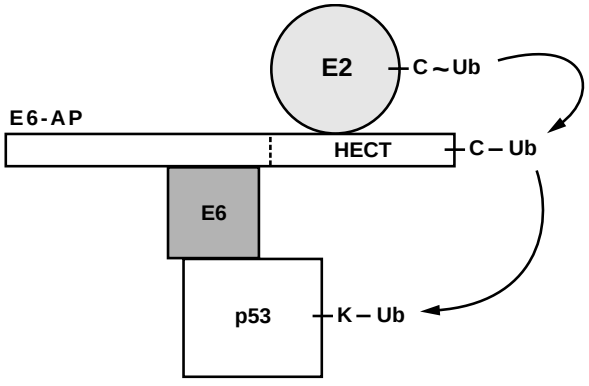


Fig. 2.11 Ubiquitinylation of the tumor suppressor protein p53. The attachment of ubiquitin molecules to p53 is catalyzed by the Hect-domain E3 ubiquitin ligase E6-AP which receives the activated ubiquitin from an E2 enzyme. An intermediate is formed where the ubiquitin is linked via a thioester bond to an active site cysteine of E6-AP. The E6 protein binds both to E6-AP and p53 and confers substrate specificity. C. cysteine; K: lysine.

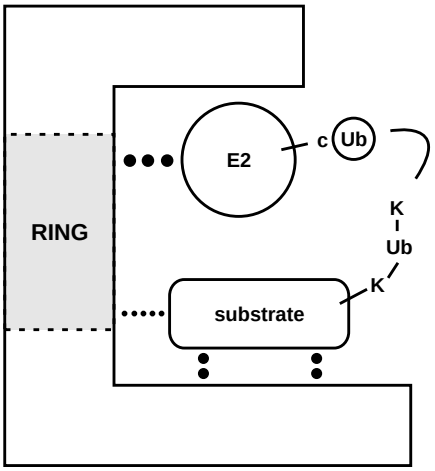


Fig. 2.12 Model of function of RING finger E3 enzymes. The E3 ligase is shown as a continuous block with three putative functional domains interacting with protein substrate and an E2. The RING finger of E3 enzymes has been defined as a primary binding site for E2. The primary binding site for the substrate may be provided by a domain containing a F-box motif. This might be a separate polypeptide from the RING finger. A secondary binding site for the protein substrate may exist in the RING finger. The E2 active-site cysteine (C) is shown forming a thioester bond with ubiquitin which is subsequently transferred to lysine residues (K) involved in formation of a polyubiquitin chain.

Interestingly, the RING finger motif is found in a large number of proteins of recognized regulatory function that hitherto have not been associated with ubiquitin-proteasome-mediated proteolysis. Examples are the breast cancer susceptability gene product BRCA1 and the MDM2 protein (see Chapter 14).

It is becoming increasingly clear that RING finger-containing E3 enzymes play pivotal roles in central cellular processes. Functions of RING finger E3 enzymes in cell cycle regulation, proliferation and apoptosis have been established. In addition, they participate in endocytosis and secretion.

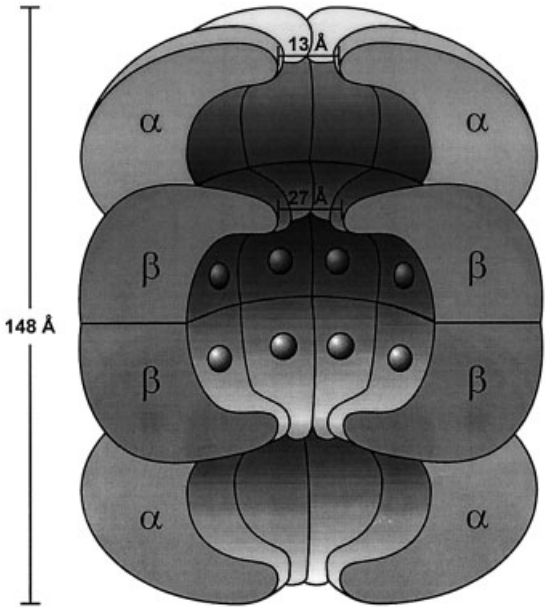


Fig. 2.13 Structure of the 20S proteosome of Thermoplasm acidophilum. The figure shows the schematic structure of the 20S proteosome (Loewe et al., 1995). Four stacked rings can be identified in the 20S proteosome, each consisting of 7 protomers. The two external rings contain 7 copies of the 26 kDa α-subunit of the proteosome, while the inner rings are composed of 7 copies of the 22 kDa β-subunit. The rings form a central channel with three chambers. The catalytic centers of proteolytic cleavage are localized on the β-subunits of the inner chamber and are represented in the figure as spheres.

2.6.2 Degradation in the Proteasome

The degradation of protein-ubiquitin conjugates occurs in an ATP-dependent reaction within a large protease complex, the *26S proteasome* (review: Baumeister et al., 1998). The substrate protein is degraded to peptides in the 26S proteasome, while the ubiquitin is released and again available to form protein conjugates. A major prerequisite for proteasomal degradation is the presence of poly-ubiquitin chains, linked by K48 isopeptide bonds, on the substrate protein. Assembly of poly-ubiquitin chains through other Lys residues like, e.g., K 63 seems to serve other functions and apparently directs the protein substrates to other routes, like DNA repair, stress response and endocytosis (review: Pickart, 2000).

The 26S proteasome is composed of two protein aggregates, a 19S and a 20S particle. The main proteolytic component of the 26S proteasome is the 20S particle, the structure of which from an archaebacterial system and a yeast system has been solved (Löwe et al., 1995; Groll et al., 1997).

The structure of the 20S proteasome (Fig. 2.13) from *Thermoplasma acidophilum* displays four rings stacked upon each other surrounding a central cavity in which proteolysis takes place. An N-terminal threonine has been identified as an essential active site residue of the protease center. The OH-group of the threonine functions as a nucleophile during hydrolysis of the petide bond. A similar mechanism of hydrolysis has been shown for other hydrolases, which, because of this property, are now included in the family of N-terminal nucleophile hydrolases. For some β-subunits of eucaryotes the N-terminal threonine is generated by autoproteolysis of an N-terminal prosequence.

The presence of the protease center in the central cavity ensures that the proteolysis is compartmentalized and shielded from the surrounding media. The substrate proteins are accessible to the proteolytic center only via a ring-shaped opening at the end of the 20S proteasome and require the assistance of the 19S particle. The structure of the 20S proteasome also indicates that proteins are accessible to the catalytic center only in the unfolded state.

The 20S proteasome from yeast and from vertebrates (review: Orlowski and Wilk, 2000) has a similar overall structure to that of the archaebacterial proteasome, but shows a more complicated subunit structure. It is functionally more versatile, and distinct variants of it can be used in specific degradation reactions. The yeast 20S proteasome is organized in four stacked rings, each containing seven subunits surrounding a central tunnel. Each of the two outer rings contains seven different α-subunits, whereas each of the two inner rings is built up from seven different β-subunits. The subunits occupy specific positions in the 20S proteasome. In the vertebrate 20S proteasome, three distinct catalytic centers have been identified on specific β-subunits, designated X, Y and Z. The X-subunit expresses a chymotrypsin-like activity, the Y subunit a peptidyl-glutamyl-peptide-hydrolyzing activity and the Z subunit a trypsin-like activity.

A distinct feature of the vertebrate proteasome is the possibility to exchange specific subunits thereby generating proteasome variants. By exchange of β-subunits different

20S complexes can form which function in specific degradation reactions. In the proteasome that is involved in processing antigenic peptides the catalytic reactions are performed by ß-subunits different from the the X,Y and Z subunits. This specific proteasome is also called 'immunoproteasome'.

The binding of ubiquitinylated substrates requires the 19S complex of the proteasome, which possesses a ubiquitin binding site and several ATPase sites. It is believed that the function of the 19S particle is to recognize the ubiquitinylated protein, unfold the protein, remove the ubiquitin moiety, and transport the unfolded protein into the catalytic chamber of the 20S proteasome.

2.6.3 Recognition of the Substrate in the Ubiquitin-Proteasome Degradation Pathway

The ubiquitin-proteasome pathway is a central pathway for the turnover of short-lived proteins. Its substrates include numerous regulatory proteins whose turnover is frequently an integral part of their function in regulation and signal transduction. It is therefore evident that the target proteins must exhibit unique features in order to ensure specific ubiquitin conjugation. It is now generally assumed that the E3 ubiquitin ligases are responsible for the selection of proteins for ubiquitinylation and subsequent degradation. Because of the large variety and incomplete characterization of E3 enzymes a commonly accepted classification is not yet available. According to the nature of the target sequences recognized and the regulation of their function the following types of E3 enzymes will be highlighted (Hershko and Ciechanover, 1998).

N-end Rule Enzymes

This class of E3 enzymes recognizes the nature of the N-terminus of target proteins. As shown by Varshavsky (review: Varshavsky, 1996), yeast proteins may be selected for ubiquitinylation according to the "N-end rule". Proteins containing basic or bulky hydrophobic amino acids at the N-terminus are recognized by distinct N-end rule E3 enzymes of which the E3*α* enzyme (UBR1 protein in yeast) is best characterized.

Cyclosome, Anaphase-promoting Complex (APC)

The cyclosome (or APC) is a high-molecular-weight complex of at least 12 different subunits that degrades proteins containing a specific recognition sequence, the destruction box (see Section 13.2.4). The E3 activity has been shown to reside in a small RING finger protein of the APC, Apc11p.

Substrates are cell cycle regulators as, e.g., cyclins, kinase inhibitors and spindle-associated proteins. Importantly, some forms of the cyclosome require phosphorylation in order to be active.

Cbl Protein

The Cbl proteins are single-subunit E3 ubiquitin ligases that are involved in the down-regulation of transmembrane receptors. They interact with phosphotyrosine residues on activated receptor tyrosine kinases via a specific SH2 domain (see Section 8.2.1),

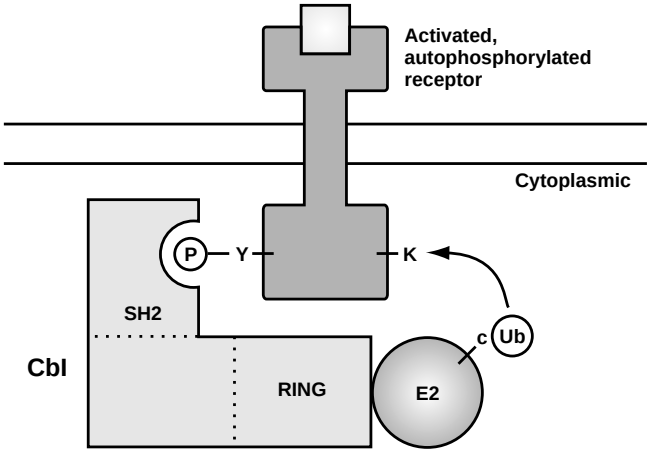


Fig. 2.14 Cbl-induced ubiquitinylation of transmembrane receptors. Ligand binding to receptor tyrosine kinases triggers autophosphorylation on the cytoplasmic region of the receptor. The tyrosine-phosphates are bound by an SH2 domain of the Cbl protein which is a E3 ubiquitin ligase. The RING finger motiv of Cbl mediates binding of an E2 enzyme from which ubiquitin is transferred to acceptor lysine residues of the receptor inducing its proteasomal degradation.

also called tyrosine kinase binding domain and they contain a RING finger domain that mediates the binding of E2 enzymes (Zheng et al., 2000). In the complex formed a transfer of ubiquitin to the receptor occurs, and the receptor is thereby targeted for endocytosis and also for degradation (Fig. 2.14).

The signal that induces the ubiquitination is the autophosphorylation of the cytoplasmic side of the receptor that takes place as a consequence of binding of the extracellular ligand to the receptor. By this mechanism the concentration of activated receptors at the membrane is reduced and signal transduction is weakened.

The signaling molecules that are downregulated by the Cbl proteins include receptor tyrosine kinases like the epidermal growth factor receptor (EGFR, see Chapter 8) and nonreceptor tyrosine kinases like Zap 70 (see Section 11.3). The central importance of the Cbl proteins for cellular regulation is highlighted by the observation that oncogenic forms of Cbl have been found in mouse retroviruses.

Phosphoprotein-Ubiquitin Ligase Complexes

Another type of multisubunit ubiquitin ligase, the SCF complex, is involved in the degradation of cell cycle regulators such as the CDK inhibitor Sic1 and G1 cyclins (see Section 13.2.4). It is characteristic of this class of ubiquitin ligases that the substrates must be phosphorylated in order to become a substrate for the ubiquitin ligase complex. Protein degradation in this system can be highly regulated since signal path-mediated phosphorylation can convert a substrate into an active form that is susceptible to ubiquitinylation by the ligase complex. Very often the regulatory phosphorylation events are found in sequence elements called PEST sequences (see Section 13.3.1).

2.6.4 Regulatory Function of Ubiquitin Conjugation and the Targeted Degradation of Proteins

The normal functioning and growth of a cell requires that the life span, and thus the activity, of regulatory proteins and key enzymes be limited. To this end, ubiquitin conjugation and the subsequent proteolysis in the proteasome represents an important tool.

Ubiquitin conjugation has been correlated with the following cellular functions:

- degradation of proteins under stress situations
- degradation of denatured and damaged proteins
- targeted degradation of regulatory proteins: oncoproteins, tumor suppressor proteins
- transmembrane receptors, mitotic cyclins, transcription activating proteins
- modulation of the activity of cell surface receptors
- protein import into cellular organelles
- repair of DNA
- processing and presentation of antigens
- assembly of ribosomes.

Of particular interest for regulatory processes are mechanisms by which the activity of growth-regulating proteins and central transcription factors are controlled via ubiquitinylation. Often the cell uses signal path-mediated protein phosphorylation in order to induce the regulated degradation of a signal protein. Examples are the G1 cyclins, the tumor suppressor p53 and the inhibitor IκB.

Two examples will be discussed in more detail.

Tumor Suppressor Protein p53

The p53 protein assumes an important role in the control of growth of higher organisms. It functions as a tumor suppressor, i.e., it suppresses the growth of tumors (see Chapter 14). An inactive, mutant form of p53 is found in nearly half of all human tumors.

Changes in the concentration of p53 induced by the ubiquitin-proteasome pathway are of central importance in the p53 regulatory network. p53 can be degraded by at least two ubiquitin-dependent mechanisms. One degradation pathway involves the Mdm2 protein (see Section 14.8.6), and the other is linked to infection by oncogenic DNA viruses as, e.g., the human papilloma virus. The viral protein involved is the oncoprotein E6 of human papilloma virus and a cellular E3 enzyme, termed E6-AP (E6-associated protein), which belongs to the Hect familiy of E3 enzymes. The ubiquitination of p53 then leads to its degradation, resulting in the loss of an important growth control in the cell. Recognition of p53 and transfer of ubiquitin occurs in a complex between the viral E6 protein, E6-AP and p53 (see Fig. 2.11). In this reaction, p53 is first bound by the E6 protein, and then recognizes the p53-bound E6 protein and transfers the ubiquitin in collaboration with the appropriate E2 enzyme to an acceptor lysine on p53. The reaction involves a covalent intermediate where the ubiquitin is linked to an active site cysteine of E6-AP. The ubiquitinylation initiated by the E6 protein and the

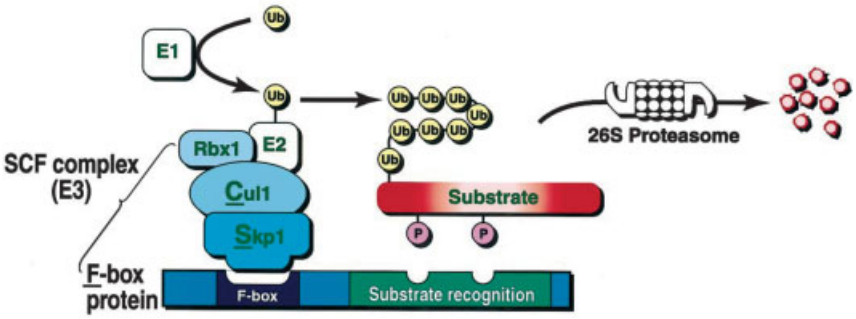


Fig. 2.15 Structure and function of the SCF complex. The SCF complex was originally identified as a heterotrimer composed of Skp1, Cul1 and an F-box protein. A fourth component, the RING finger protein Rbx1, was subsequently identified. Each F-box protein contaions an F-box domain, which is required for binding to Skp1, as well as a specific substrate-recognition motif (from Nakayama et al., 2001).

ensuing degradation of p53 result in a loss of p53 function, thus offering an explanation for the tumor-causing effect of the papilloma virus.

NFκB

The transcription activator NFκB regulates a variety of genes involved in the immune response and the inflamatory process (review: Perkins, 2000). NFκB is required for the expression of genes for the light x-chain of immunoglobulins, for interleukin 2 and 6, and for interferon b (see Chapter 11).

The function and regulation of NFκB is shown schematically in Fig. 2.16. The active form of NFκB is a heterodimer consisting of one p50 and one p65 subunit. In the cytosol NFκB is found in an inactive complex bound to the inhibitor protein IκB. IκB masks the nuclear translocation signal of the heterodimer, thus preventing its transport into the nucleus. The activity of NFκB is highly regulated.

NFκB is activated upon extracellular stimuli, as, e.g., the action of growth factors, cytokines (see Chapter 11) or the exposure to UV light. The signal pathway that leads to phosphorylation and subsequent degradation of NFκB has been well characterized for the cytokines IL-1 and for Tumor Necrosis Factor, TNF (review: Chen and Goeddel, 2002). Following binding of a cytokine to its transmembrane receptor, a family of specific protein kinases including a high molecular mass IκB kinase complex is activated that phosphorylate the inhibitor IκB. This phosphorylation is the signal for ubiquitinylation and degradation of IκB. NFκB is thus released from its inhibited state to translocate in the nucleus and activate transcription of target genes.

The ubiquitin-proteasome pathway participates in the regulation of NFκB at two points:

- The p50 subunit of NFκB results from the proteolytic processing of a 105 kDa precursor protein (p105) in the cytosol. The processing requires the poly-ubiquitinylation of p105 mediated by the 26S proteasome.
- The degradation of the inhibitor protein IκB involves the ubiquitin-proteasome pathway.

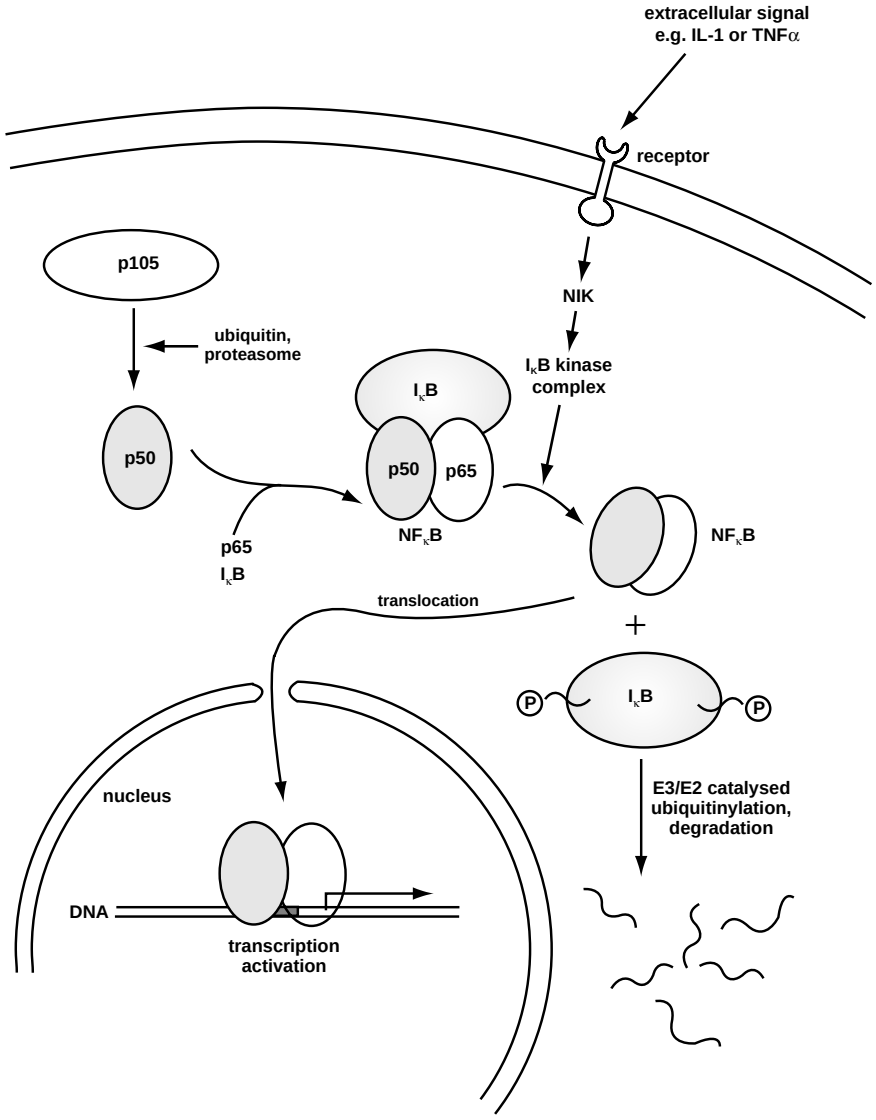


Fig. 2.16 Regulation and proteolysis of the transcription factor NFκB. The Ubiquitin-proteasome pathway is involved in the regulation of NFκB in two ways. The 50 kDa subunit of NFκB is formed from a 105 kDa precursor via ubiquitin mediated proteolysis. In the cytosol NFκB is heterodimeric and inactive, bound to the inhibitor protein IkB. Upon an external signal, the phosphorylation of NFkB is induced, whereby the inhibitor complex is broken up. The inhibitor IκB is degraded via the ubiquitin-proteasome pathway, while the phosphorylated NFκB is transported to the nucleus where it can fulfill its transcription regulation function.

The ubiquitin-proteasome system thus has significance for NFκB in two ways. On the one hand, it participates in the specific processing of the p105 precursor protein to the small subunit of NFκB. On the other hand, NFκB is activated because of the degradation of IκB.

This example nicely illustrates how extracellular signals can induce the ubiquitinylation and degradation of specific proteins. As shown by the processing of the p105 precursor, ubiquitinylation can also be used for partial proteolysis and for specific activation of a regulatory protein.

2.7 Regulation of Proteins by Sumoylation

Sumoylation is a covalent modification of proteins that is related to, but functionally distinct from ubiquitination (review: Wilson and Rangasami, 2001). As in ubiquitinylation, sumoylation involves the covalent attachment of a small protein moiety, termed SUMO, to target proteins. The reactions leading to sumoylation of substrate proteins are related to those involved in ubiquitination. E1- and E2 like enzymes are responsible for the attachment of the SUMO moiety to lysine residues of the target protein. As compared to ubiquitination, sumolyation is more sequence specific and requires a particular amino acids in the neighbourhood of the lysine to be modified.

The functions of sumolyation appear to be more diverse than that of ubiquitination and the mechanistic basis of control by sumoylation is incompletely understood. As an example, SUMO modification of RanGAP1 (see Section 9.1 5.) regulates its subcelluar localization by targeting it to the nuclear pore complex. SUMO modification of the inhibitor IκB (see above, Fig. 2.16) occurrs at the same Lys-residues that are used for attachment of ubiquitin molecules, thus precluding ubiqutination. As a result, IκB is stabilized and remains bound to NFκB inhibiting nuclear transclocation of NFκB and transcriptional activation. By contrast, the tumor suppressor p53 is modified in the same region by SUMO and ubiquitin, but the two processes do not compete. Ubiquitination of p53 targets it for proteolysis whereas SUMO modification increases its transcriptional activity.

Reference

1 Barford, D., Hu, S.H. and Johnson. L.N. 'Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP' (1991) *J. Mol. Biol.* **218**, 233–260

2 Baumeister, W., Walz, J., Zuhl, F. and Seemuller, E. (1998) The proteasome: paradigm of a self-compartmentalizing protease. *Cell*, **92**, 367–380.

3 Chen, G. and Goeddel, D.V. (2002) TNF-R1 signaling: a beautiful pathway. *Science*, **296**, 1634–1635.

4 Fersht, A. 'Enzyme Structure and Mechanism' (1998) Freeman, New York

5 Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H.D. and Huber, R. (1997) Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature*, **386**, 463–471.

6 Hershko, A. and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.*, **67**, 425–479.

7 Hurley, J.H., Dean, A.M., Sohl, J.L., Koshland, D.E. and Stroud, R.M. (1990) 'Regulation of an enzyme by phosphorylation at the active site' *Science* **249**, 1012–1016

8 Joazeiro, C.A. and Weissman, A.M. (2000) RING finger proteins: mediators of ubiquitin ligase activity. *Cell*, **102**, 549–552.

9 Johnson L.N. and O'Reilly M. 'Control by phosphorylation' (1996) Curr Opin Struct Biol. 6, 762–769

10 Kraut, J. (1988) 'How do enzymes work?' *Science* **242**, 533–540

11 Krebs, E.G., Graves, D.J. and Fischer, E.H. (1959) 'Factors affecting the activity of muscle phosphorylase kinase' *J. Biol. Chem.* **234**, 2867–2873

12 Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. and Huber, R. (1995) Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. *Science*, **268**, 533–539.

13 Nakayama, K.I., Hatakeyama, S. and Nakayama, K. (2001) Regulation of the cell cycle at the G1-S transition by proteolysis of cyclin E and p27Kip1. *Biochem. Biophys. Res. Commun.*, **282**, 853–860.

14 Orlowski, M. and Wilk, S. (2000) Catalytic activities of the 20 S proteasome, a multicatalytic proteinase complex. *Arch.Biochem.Biophys.*, **383**, 1–16.

15 Perkins, N.D. (2000) The Rel/NF-kappa B family: friend and foe. *Trends Biochem.Sci.*, **25**, 434–440.

16 Pickart, C.M. (2001) Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.*, **70**, 503–533.

17 Varshavsky, A. (1996) The N-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci. U.S.A*, **93**, 12142–12149.

18 Wilson,V.G. and Rangasamy, D. (2001) Intracellular targeting of proteins by sumoylation. *Exp. Cell Res.*, **271**, 57–65.

19 Zheng, N., Wang, P., Jeffrey, P.D. and Pavletich, N.P. (2000) Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell*, **102**, 533–539.

3
Structure and Function of Signal Pathways

3.1
General Function of Signal Pathways

The enormous structural variety and functional capacity of multicellular organisms is due to their ability to coordinate the biochemical reactions of the various cells of the total organism. The basis for this coordination is the intercellular communication, which allows a single cell to influence the behavior of other cells in a specific manner.

We currently know of various forms of communication between cells (Fig. 3.1):

1. *Chemical Messengers*: Cells send out signals in the form of specific chemical messengers that the target cell transmits into a biochemical reaction. Signaling cells can simultaneously influence many cells by chemical messengers so as to enable a temporally coordinated reaction in an organism.
2. *Gap Junctions*: Communication between bordering cells is possible via direct contact in the form of "gap junctions". Gap junctions are channels that connect two neighboring cells to allow a direct exchange of metabolites and signaling molecules between the cells.
3. *Cell-cell interaction via cell surface proteins*: Another form of direct communication between cells occurs with the help of surface proteins. In this process a cell surface protein of one cell binds a specific complementary protein on another cell. As a consequence of the complex formation, an intracellular signal chain is activated which initiates specific biochemical reactions in the participating cells.

A further intercellular communication mechanism relies on electrical processes. The conduction of electrical impulses by nerve cells is based on changes in the membrane potential. The nerve cell uses these changes to communicate with other cells at specialized nerve endings, the synapses. It is central to this type of intercellular communication that electrical signals can be transformed into chemical signals. This type of communication will not be discussed in this book.

Intercellular signal transduction influences nearly every physiological reaction. It ensures that all cells of a particular type receive and transform a signal. In this manner, cells of the same type react synchronously to a signal. A further function of signaling pathways is the coordination of metabolite fluxes between cells of various tissues.

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

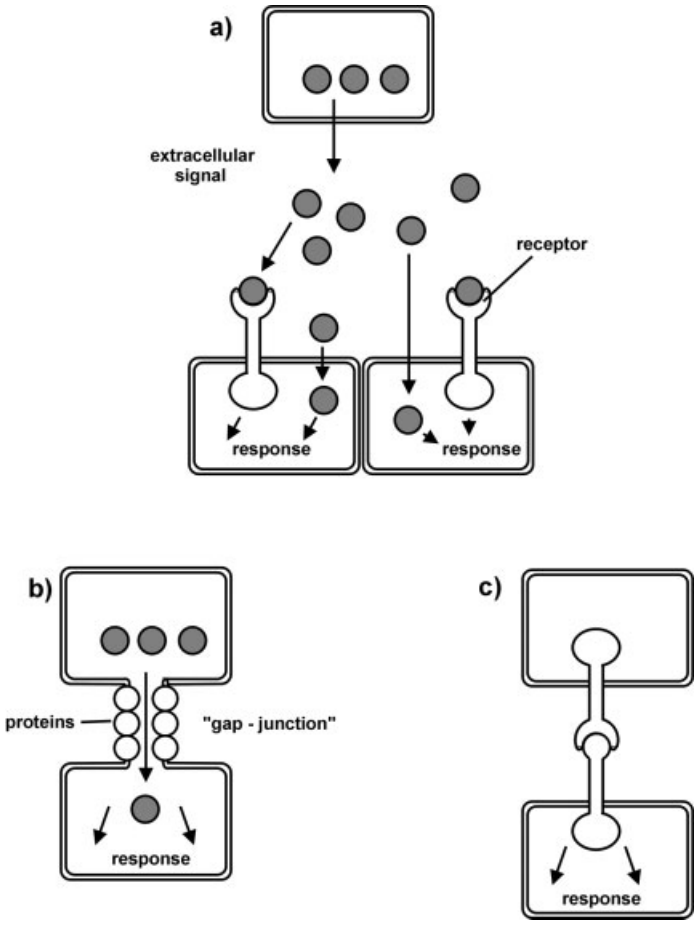


Fig. 3.1 Principal mechanisms of intercellular communication. a) communication via intercellular messengers. b) communication via gap junctions. Gap junctions are direct connections between cells. They are coated by proteins (drawn as circles in the figure above) that can have a regulatory influence on the transport. c) communtication via surface proteins.

In higher organisms intercellular signaling pathways have the important task of *coordinating and regulating cell division*. The pathways ensure that cells divide synchronously and, if necessary, arrest cell division and enter a resting state.

Cellular communication assumes great importance in *the differentiation and development* of an organism. The development of an organism is based on genetic programs that always utilize inter- and intracellular signaling pathways. Signal molecules produced by one cell influence and change the function and morphology of other cells in the organism.

Signaling pathways are also critical for the processing of *sensory information*. External stimuli, such as optical and acoustic signals, stress, gradients of nutrients, and so on,

are registered in sensory cells and are transmitted to other cells of the organism via signaling pathways.

3.2 Structure of Signaling Pathways

Intercellular communication relies on the creation of specific signals by a signaling cell. The signals are registered by a target cell and are thereupon transmitted and processed further with the help of intracellular signal chains.

3.2.1 The Mechanisms of Intercellular Communication

In the communication between cells of an organism, the signals (chemical messengers or electrical signals) are produced in specialized cells. The signal-producing function of these cells is itself regulated, so that the signal is only produced upon a particular stimulus. In this way signaling pathways can be coupled to one another and coordinated.

The following steps are involved in intercellular communication (Fig. 3.2):
- formation of a signal in the signal-producing cell as a result of an external trigger
- transport of the signal to the target cell
- registration of the signal in the target cell
- further transmission of the signal into the target cell
- transformation of the signal into a biochemical or electrical reaction in the target cell
- termination of the signal.

A target cell that receives a signal within the framework of intercellular communication transmits the signal in *intracellular* pathways. The intracellular signaling pathways are characterized by the following parameters:
- nature of the triggering, external signal
- mechanism of the registration of the signal
- mechanism of the transmission and termination of the signal
- nature of the biochemical reaction induced in the target cell.

The sum of these reactions determines the response of the target cell.

Nature of the External Signal

Cells can receive and process signals in the form of messenger substances (proteins, low-molecular-weight substances), and electrical, optical and other stimuli.

Reception of Signals by Receptors

Specialized proteins, termed receptors, are utilized for the reception of signals. The reception of the signals by the receptor is equivalent to the binding of messenger

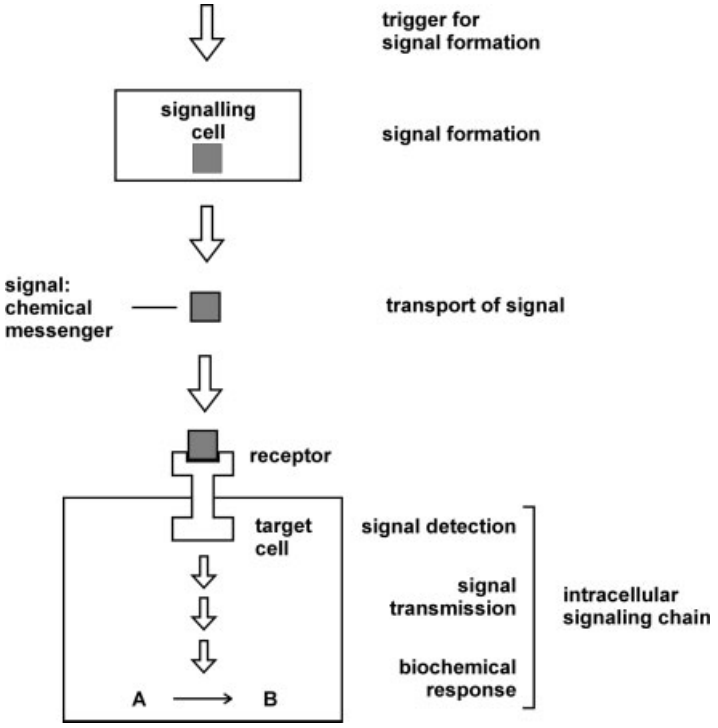


Fig. 3.2 The individual steps of intercellular communication. Upon reception of a triggering stimulus, the signal is transformed into a chemical messenger within the signaling cell. The messenger is secreted and transported to the target cell, where the signal is registered, transmitted further, and finally converted into a biochemical reaction. Not shown are processes of termination or regulation of communication which can act at any of the above steps.

substance on the receptor or the transmission of physical stimuli into a structural change in the receptor.

There are two principal ways by which target cells can process incoming signals:

- Cell surface receptors receive the signal (e.g., a messenger substance) at the outside of the cell, become activated and initiate a signaling chain in the interior of the cell. In such signaling pathways the membrane-bound receptor transduces the signal at the cell membrane so that it is not necessary for the signal to actually enter the cell.
- The messenger enters into the target cell and binds and activates the receptor localized in the cytosol or nucleus.

Transmission of the Signal

Upon receiving a signal, a receptor becomes activated to transmit the signal further. The *activated receptor* passes the signal onto components, usually proteins, further downstream in the signaling pathway, which then become activated themselves for

further signal transmission. A chain of serially operating, intracellular signal transduction processes results. Finally, a specific biochemical process is triggered in the cell, which represents the endpoint of the signaling pathway.

Regulation and Termination

Cells possess multiple mechanisms to regulate the intercellular communication as well as the intracellular signal transduction. This allows a specific attenuation or termination of communication between cells. Often feedback mechanisms are used to adopt the cellular response to the needs of the organism.

3.2.2 Principles of Intracellular Signal Transduction

In the following we shall deal with the basic components and principles of intracellular signal transduction and signal processing. The specific reactions and the various levels of signal transduction will be discussed in detail in later chapters.

Components of the intracellular signal transduction are proteins and small molecule messengers (Fig. 3.3). An incoming signal is passed on from the receptor to downstream proteins, which themselves have other proteins as the next partner in the signal cascade. In this manner, further signaling proteins are recruited to act in the signaling chain. The participating signaling proteins can be *enzymes* or they can act as connectors to recruit the other proteins in the signaling pathway. The latter type of proteins are termed *adaptor proteins*.

The transduction of a signal in a signaling chain occurs by a sequential activation of the participating proteins. Important mechanisms for the activation of signaling proteins are

- chemical modification of signaling proteins
- triggering of enzymatic activity
- triggering of conformational changes
- increase of the concentration of signaling proteins
- colocalization of signaling proteins.

Very often the proteins of a signaling chain are multivalent, i.e. they can receive signals from more than one upstream signaling partner and they can pass on the signal to more than one downstream effector proteins. This property allows a branching of signaling pathways and the formation of signaling networks.

Another important principle of signaling transduction is the use of diffusible chemical messengers that are formed upon the activation of signaling enzymes. By specifically binding to cognate signaling proteins these are activated for further signal transmission.

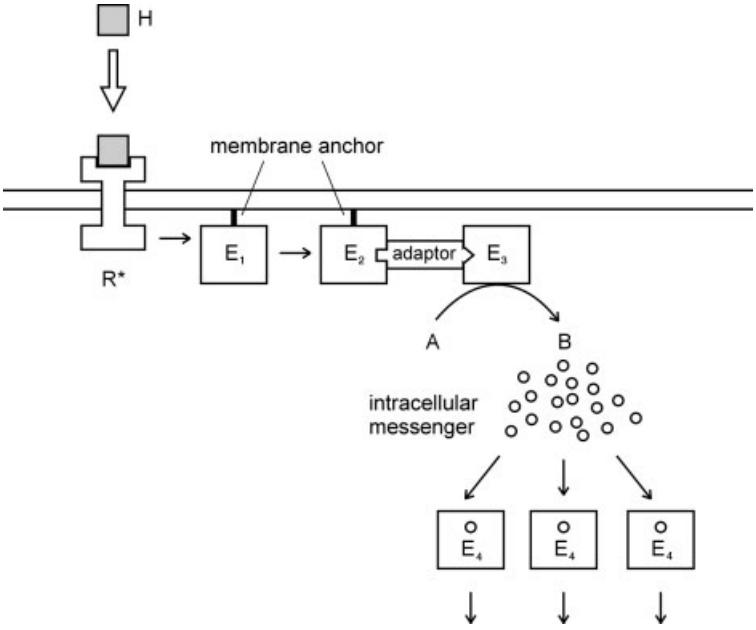


Fig. 3.3 Components of intracellular signal transduction. The reception of an extracellular signal by a membrane receptor, shown here as binding of a hormone to its receptor, activates the receptor for further signal transduction. The activated receptor R* passes the signal onto downstream effector proteins E. Adaptor proteins may be involved in the pathways between effector proteins. The transduction of signal from the receptor to its downstream effector is usually a membrane-associated process. The example shown in the diagram above is only to be construed as an example for the composition of a generic signaling pathway. The structure of the intracellular signaling pathways of a cell are highly variable. There are signal transduction pathways that are much simpler than the one represented in the figure above, and others that involve many more components and are much more complicated.

3.2.3

Components of Intracellular Signal Transduction

Signaling Proteins

The most important tools for intracellular signal transduction are the receptors, protein kinases, protein phosphatases, regulatory GTPases and adaptor proteins:

- **Receptors**
  Receptors are the gates for incoming signals in that they specifically receive the signal and are thereby activated for further signal transduction. The main properties of receptors will be discussed below (see Section 3.4)
- **Protein kinases and protein phosphatases**
  A central tool for signal transmission in a cell is the phosphorylation of proteins via protein kinases. Proteins can be reversibly activated or inactivated via phosphorylation. The phosphorylation status of a protein is controlled by the activity of both

protein kinases and protein phosphatases (see Chapter 7). Both classes of enzymes are elementary components of signaling pathways and their activity is subject to manifold regulation.

- **Regulatory GTPases**
  The regulatory GTPases (see Chapter 5) function as switches that can exist in an active or inactive form. In the active form the GTPases can transmit signals to downstream components in the signaling chain. In the inactive form signal transmission in repressed.
- **Adaptor proteins**
  Adaptor proteins (see Chapter 8) mediate the signal transmission between proteins of a signaling chain by bringing these proteins together. They function as clamps to co-localize proteins for an effective and specific signaling. Furthermore, adaptor proteins help to target signaling proteins to specific subcellular sites and to recruit signaling molecules into multiprotein signaling complexes. In the latter case, the adaptor proteins may function as a scaffold or docking site for organizing different signaling molecules at distinct sites. The proteins are then also termed docking or scaffolding proteins.

Diffusible Intracellular Messengers

The intracellular activation of enzymes in a signaling chain can lead to the formation of diffusible chemical signaling molecules in the cell. These intracellular signaling molecules are also termed *second messengers* (see Chapter 6). The second messenger molecules activate and recruit cognate enzymes for the further signal transduction.

The following properties are important for the function of diffusible intracellular messengers:

The second messengers
– may be rapidly formed from precursors by enzymatic reactions
– may be rapidly released from intracellular stores
– may be rapidly inactivated or stored in specific compartments
– may activate different effector proteins
– allow the amplification of signals.

Contrary to what is suggested by the term 'diffusible intracellular messengers', these signaling molecules normally do not diffuse across the whole cytoplasmic space. Rather, the second messengers are often used to create signals that are limited in time and in space. The second messengers can be formed as well as inactivated in specific compartments and at specific sites of the cell membrane resulting in locally and timely restricted reactions. This is mainly true for the second messengers Ca^{2+} and cAMP that are formed or released only at specific sites and often function only locally (see Chapter 6). Another important property of the second messengers is their ability to activate very different effector proteins. As an example, Ca^{2+} binds and activates a large number of different effector proteins. The release of Ca^{2+} induced by an external stiumulus thus can trigger very different downstream biochemical reactions (see Section 6.7).

3.2.4 Coupling of Proteins in Signaling Chains

In order to transmit a signal into a defined biochemical answer, the various components in a signaling chain must interact in an unequivocal, specific way. Cells have available a broad repertoire of mechanisms that ensure a specific coupling of signaling proteins and prevent both undesired side reactions and the flowing away of the information that is contained in a triggering signal. The coupling mechanisms enable the cell to guide the information in a precise way to the biochemical end points of the resulting signal transduction chain. By this token, both the nature of the signaling components and the mechanisms by which these components interact are essential to the structure of a signaling pathway.

Another essential feature of signaling pathways is the ability to regulate the coupling of the signaling proteins. The interaction between signaling proteins can be activated or inactivated, which allows signals to be active in a timely and locally defined way. This is achieved by, e.g., conformational changes, the reversible creation of docking sites, and changes in the subcellular localization of the signaling proteins.

3.2.4.1 Coupling by Specific Protein–Protein Interactions

The interaction and thus coupling between signaling proteins can occur via structural domains that are characteristic for a given signaling protein and are more or less of a singular character. An example is the effector domain of the Ras-protein (see Chapter 9) that mediates the interaction with a complementary binding region on the downstrem effector protein, the Raf kinase. In addition to the Raf kinase, other downstream effectors can interact with the Ras effector domain as well, albeit with different affinities. Importantly, the availability of the binding surface for interaction with the downstream effector can be regulated by allosteric mechanisms. In this way, signaling proteins can exist in an activated on-state where the binding surface is exposed or an inactivated off-state, where the binding surface is buried and signal transduction is not possible.

3.2.4.2 Coupling by Protein Modules

A common principle of protein–protein interactions in signaling chains uses protein modules that are found in many signaling proteins with nearly identical or related structure and function. The binding partners of these protein modules are often small sequence motifs or modified amino acids that are found on the cognate effector proteins. The function of these protein modules will be dealt with in detail in Chapter 8.

Based on a similar basic structure, many variations of the binding modules are found in signaling proteins that differ in the exact nature of the complementary binding element and thus have distinct sets of effector proteins as binding partners. The use of a similar binding module in different signaling proteins is certainly a very economic way to mediate protein-protein interactions. Nevertheless, the specificity of interaction is ensured since the various variants of a given protein module have different binding preferences.

Signaling proteins often contain several protein modules that mediate the coupling to different binding partners. These signaling proteins are multivalent and can mediate the clustering of different proteins into larger aggregates where signal transduction can take place in an efficient way. Moreover, the presence of several signaling modules in a given protein allows the participation in several distinct signaling pathways. For signaling proteins harboring many protein modules, e.g., the PDGF-receptor (see Section 8.2.1 and Fig. 8.9), it is improbable that all possible binding partners will dock onto the protein at the same time. Rather, the presence of many modules offers the possibilty to use each one in dependence on the availability of the binding partners, which provides an increased flexibiltiy and regulation of signaling.

3.2.4.3 **Coupling by Reversible Docking Sites**

A central tool for the coupling of proteins in signaling chains is the phosphorylation of tyrosine and – to a lesser extent – Ser/Thr residues (see Chapter 2). By this chemical modification, a reversible docking site is created for effector proteins that contain protein modules (SH2, PTB domains, see Section 8.2) with binding specificity for the phosphorylated site. This coupling mechanism is regulated by the activity of protein kinases and protein phosphatases that are themselves components of the same or another signaling chain. Protein tyrosine kinases activate the coupling mechanism, whereas it is inactivated by the tyrosine phosphatases. The intensity of signaling at this point depends on the relative activites of the two enzyme species. Both activities can be actively regulated during signaling.

The coupling of two signaling proteins by means of Tyr-phsophorylation and SH2 domains is a very versatile mechanism. Many signaling proteins contain multiple Tyr-phosphorylation sites, each in a different sequence context. On the other hand, there are many different SH2 domains, each with preference for distinct sequence context of the Tyr-P. A signaling protein harboring many Tyr-phosphorylation sites may therefore interact with different SH2 domains found on very different signaling proteins. A prominent example is the PDGF receptor whose multiple tyr-phosphorylation sites allow – at least in theory – the docking of eight different effector proteins

3.2.4.4 **Coupling by Colocalization**

The cell often uses the strategy of colocalization for the efficient coupling of signaling proteins. By bringing together different signaling proteins at specific sites of the cell, the information transfer between the signaling components is strongly enhanced. This type of coupling is of outstanding importance for signaling processes at the cytoplasmic face of the cell membrane.

The colocalization of signaling proteins at the cell membrane is mediated by
- membrane association via lipid anchors (see Section 3.7)
- binding to transmembrane proteins by protein modules (see Section 8.2)
- interaction with specific localization subunits (see Section 7.7).

By bringing together two signaling proteins at the same site of the cell (e.g., in focal adhesion points of the cell membrane) an efficient communication between the pro-

teins is possible. Because of the immobilization of the proteins at the cell membrane, a high local concentration is achieved, and signals can be very efficiently transmitted within the assembled complex. Furthermore, the specificity of signaling is enhanced. Potential binding partners that are not membrane associated and are localized, e.g., in the cytosol, do not participate in signaling.

The colocalization of signaling proteins is essential for the signaling at the cell membrane. The transduction of signals by transmembrane receptors into the interior of the cell occurrs in multiprotein complexes assembled specifically at the inner face of the cell membrane. A large part of the Ras signaling pathway (see Chapter 9) is intimately linked to the cytoplasmic side of the membrane. If membrane association of the components of the Ras signaling pathway is prevented, signal transduction is shut down.

A particular advantage of the membrane association of signaling pathways is the possibility to link chemical signals with electrical signals. Electrical signals that trigger the influx of ions (e.g., Ca^{2+}) into the cytoplasm can act directly on membrane-associated signaling proteins. The influx of Ca^{2+} can activate directly membrane-associated forms of protein kinase C and can trigger a phosphorylation signal (see Chapter 7).

3.2.4.5 **Linearity, Branching and Crosstalk**

A signaling pathway is characterized by the type of signaling molecules involved and the order and mechanism of coupling of the various components (Fig. 3.4). We know of signaling pathways that include only a small number of different signaling molecules and are therefore of a simple structure. An example is the signaling by nuclear receptors and the signaling via the Jak/Stat pathway. These pathways allow a fast and direct flow of information to the biochemical endpoint (e.g., stimulation of gene expression) with few chances of side reactions and interference with other signaling pathways. On the other hand many signaling pathways are of a complex structure where many signaling proteins cooperate and where the information is guided to a multitude of biochemical endpoints. An example is the Ras signaling pathway, where different signals converge and a multitude of effector reactions are activated. It is an advantage of complex signaling pathways that more than one signal can be integrated and processed and that at each step a regulation and fine tuning is possible.

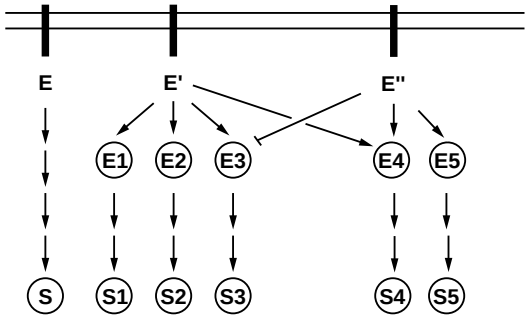


Fig. 3.4 Linearity, branching and crosstalk in intracellular signaling Crosstalk refers to a situation where a signaling enzyme from one pathway activate (E4) or inhibit (E'') signaling components involved in signal transduction of a different pathway.

Linearity of Signaling Pathways

Signaling pathways are often formulated as linear pathways where the various signaling molecules are arranged in a hierarchical order and a linear, unequivocal flow of information takes place. This is however a simplistic view of the situation in the cell and is mostly born of an incomplete knowledge of the pathway.

Branching of Signaling Pathways

As more information on the various signaling pathways of the cell was obtained the more it became clear that signaling pathways are often branched and are integrated into signaling networks. Linear pathways appear to be more the exception than the rule. Branching and network formation is mediated mainly in two ways. A given signaling molecule can be activated and can receive a signal by different upstream signaling molecules and can have different downstream effectors. Thereby, different signals can be integrated into a pathway and different signals can be generated from the pathway. Examples are the protein kinases that can be activated (or inactivated) by a multitude of mechanisms and can phosphorylate different substrates, thereby triggering a variety of downstream reactions (see Chapters 7 and 13). Furthermore, many signaling proteins are multivalent and possess binding sites for a variety of downstream effector molecules (PDGF receptor, see Fig. 8.9), allowing branching of the information transfer.

The use of second messengers is a further important tool for distributing information to different effector molecules. A second messenger like cAMP (see Chapter 6) can bind to and activate a multitude of substrates, among which the cAMP-dependent protein kinases constitute the most important and divergent class.

Signaling Networks and Crosstalk

The multivalency of signaling proteins (transmembrane receptors, adaptor proteins, Ras proteins) and the use of second messengers allows the cell to link different signaling chains and to form signaling networks. In addition, signaling proteins of one pathway (e.g., protein kinases) can covalently modify and thereby change the activity of components of other signaling pathways. Signaling in one pathway can thereby influence, regulate and modulate the signaling in other pathways. This interdependence of signaling is also called crosstalk and is presently a main topic in the research on signaling pathways.

The coupling and linkage of different signaling pathways also enables the cell to use different signals to trigger the same biochemical reaction. Furthermore, the same signal can be delivered to different pathways and thus trigger distinct reactions.

Network formation and crosstalk endows the cell with a great plasticity and variability in signal transmission. Each cell receives at a time a large number of signals and the responses on these signals must be fine tuned and orchestrated in a way that is optimal for the growth and development of the organism.

3.2.4.6 Variability and Specificity of Receptors and Signal Responses

Hormone-receptor systems and the corresponding signaling chains display great variability and flexibility (Fig. 3.5), which is expressed by the following observations:

- For a given hormone, different receptors can exist on the same or on different cells. Thus, the same hormone can trigger very distinct reactions in different tissues. An example of such a phenomenon is adrenaline, which can initiate, on the one hand, a cAMP-mediated signal transduction and, on the other hand, an inositol triphosphate-mediated reaction (see Chapter 6).
- For a given receptor, subtypes are found which differ in their affinity to the hormone, in the nature and intensity of the reaction triggered in the cell, and in their capacity for regulation.
- The same secondary reaction can be triggered by different hormone-receptor systems and signaling pathways. This is exemplified by the release of Ca^{2+}, which can be regulated via different signaling pathways (see Chapters 5 to 7).

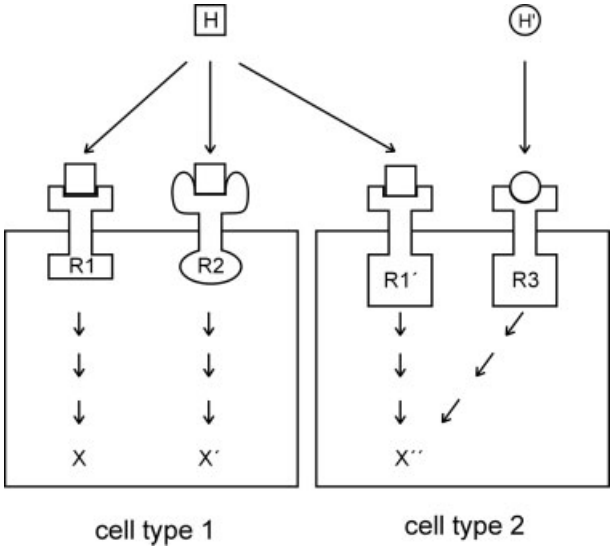


Fig. 3.5 Variability of receptor systems and signaling pathways. a) For one receptor of a given binding specificity (binding to hormone H) there can be different subtypes in the same cell (R1, R2) or in other cell types (R1'). b) The hormone H can induce different reactions (X1, X2) upon binding the different receptor types (R1, R2). The receptor types R1 and R2 can be found simultaneous in one cell. c) the binding of two different hormones (H, H') to different receptors (R1', R3) can induce the same intracellular reaction. The characteristics a) and b) contribute to a high degree to the diversity and variability of hormonal signal transduction. Point c) illustrates the principle that important cellular metabolites or reactions can be controlled by different signal transduction pathways.

Because of the above listed properties an external signal can trigger cell- and tissue-specific reactions of a highly variable nature. Experimental results on a signaling chain that have been obtained from one cell type may thus not be transferred to other cell types. Even when the overall structure of a signaling pathway in one cell type is similar to that of another cell type considerable differences may be revealed when the pathways are examined in the details of the signaling proteins, the second messengers and the coupling mechanisms of the various steps. Therefore, signaling pathways are highly cell-type specific when the details of signaling are considered.

Which Factors are Mainly Responsible for the High Variability and Cell-type Specificity of Signaling?

We now know of several of mechanisms that are used by the cell to achieve a high variability of signaling and to generate cell-type specific signaling pathways.

Variants of a signaling pathway often have a central signaling protein (e.g., Ras protein) in common that functions as a switch and can receive different signals and can transmit signals to different effector proteins. Dependent on the availability of the upstream and downstream signaling partners a variety of reactions can be triggered via the same switching station. Which signal will be guided to and transmitted by the switching station depends on the availability and activity of the upstream and downstream signaling partners. This in turn will depend in a cell-type specific way on the following points:

– cell-type specific expression of the gene for the signaling protein
– cell-type specific splicing
– cell-type specific stability
– cell-type specific post-translational modification
– cell-type specific subcellular localization

The existence of several related genes for a signaling protein, differential transcription and splicing leads to the presence of subtypes of signaling proteins that differ in the details of signal registration, signal transmission and regulation. Examples of signaling proteins for which many subtypes are found:

– α- and βγ-subunits of the heterotrimeric G-proteins (see Chapter 5)
– adenylyl cyclase (see Chapter 6)
– protein kinase C (see Section 7.4)
– protein kinase A (see Section 7.3)
– GAP proteins (see Section 9.5)
– Protein kinases of the MAPK signaling pathways (see Chapter 10)
– Jak kinases and Stat proteins (see Chapter 11)

3.3
Extracellular Signaling Molecules

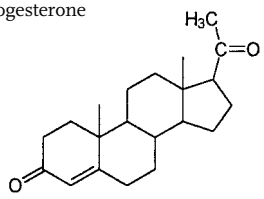
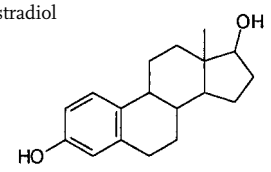
The signal-producing cells release their chemical signaling molecules either by exocytosis or passive diffusion into the extracellular space. The messengers reach their target cells via the circulatory system. In special cases, as in the communication between cells of the immune system, membrane-bound proteins are also used as signaling molecules. Communication is then only possible upon direct contact between the target cell with the surface of the signaling molecules (Fig. 3.1).

Signaling molecules for the communication between cells are known as *hormones*. Hormones that are proteins and regulate cell proliferation are known as *growth factors*.

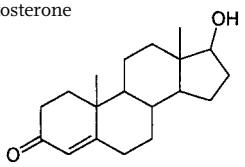
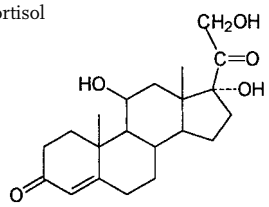
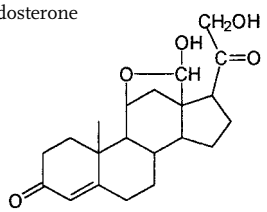
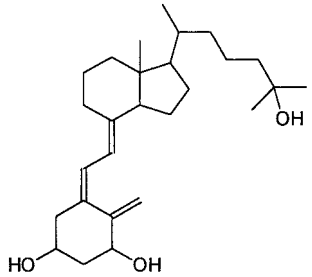
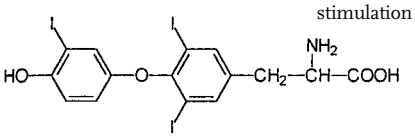
3.3.1
The Chemical Nature of Hormones

The chemical nature of hormones is extremely variable. Hormones can be:
- proteins
- peptides
- amino acids and amino acid derivatives
- derivatives of fatty acids
- nucleotides
- steroids
- retinoids
- small inorganic molecules, such as NO.

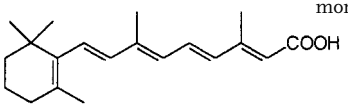
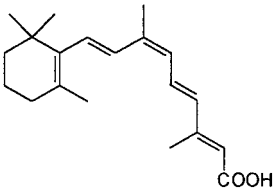
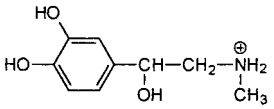
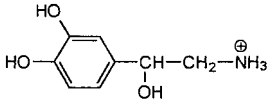
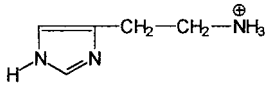
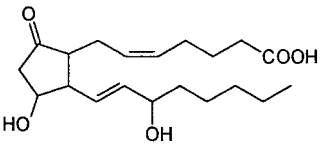
Tab. 3.1 Shows a selection of hormones from mammals and man.

a) Examples for hormones that bind to nuclear receptors	
Steroids	**biochemical and/or physiological funktion**
Progesterone 	Preparation of the uterus for implantation of the embryo, maintenance of early pregnancy
Estradiol 	Preparation of the uterus to receive the blastocyst, control of uterine constraction, generation of secretory system of breasts during pregnancy

Tab. 3.1 Continued.

Steroids	**biochemical and/or physiological funktion**
Testosterone 	Differentiation and growth of the male reproductive tract, stimulation of male secondary sex characteristics, sceletal muscle growth
Cortisol 	Metabolism of carbohydrates, lipids and proteins, anti-inflammatory, ummunsuppressive induction of Tyr-aminotransferase and of Trp-cyclooxygenase
Aldosterone 	Water and ion balance, backresorption of ions in the kidney
Steroid-relataed hormons	
1, 25-Dihydroxycholecalciferol (from vitamine D_3 	Metabolism of Ca^{2+}- and phosphate, bone mineralization, resorption of Ca^{2+} and phosphate in the intestine
Other hormones	
3,5,3'-Triiodothyronine, T_3-Hormon 	Increased oxygen consumption and increased heat formation, stimulation of glycolysis and of protein biosynthesis

Tab. 3.1 Continued.

Steroids	**biochemical and/or physiological funktion**
Retinoids	
All-trans-retinoic acid 	Formed from all-trans-retinal, broad effect on differentiation and morphogenesis
9-cis-VitA-Säure 	
b) Examples of hormones, that bind to transmembrane receptors	
Hormone	Function and biochemical action
Epinephrine 	Raise of blood pressure, contraction of smooth muscles, glycogen breakdown in liver, lipid breakdown in adipose tissue
Norepinephrine 	Contraction of arteria
Histamine 	Relaxation of blood vessels
Derivatives of arachidonic acid	
Prostaglandin E_2 	Contraction of smooth muscles

Tab. 3.1 Continued.

Peptide hormones and protein hormones	
Glucagon polypeptide, 29 aa	Glycogenolysis in liver, release of fatty acids from triglycerides in adipose tissue
Insulin polypeptid, A-chain, 21 aa, B-Kette 30 aa	Stimulation of: glucose uptake in muscle and adipose tissue, catabolism of carbohydrates, storage of triglycerides in adipose tissue, protein synthesis, cell proliferation Inhibition of glycogenolysis
Gastrin polypepetide, 17aa	HCl and pepsin in stomach
Secretion polypeptide, 27 as	Stimulation of secretion of pancreatic proteases
Adrenocorticotropin polypeptide, 39 aa	Biosynthesis in anterior pituitary, stimulation of formation of cortiosteroids in adrenal cortex, release of fatty acids from adipose tissue
Follicle stimulating hormone (FSH) polypeptide, α-chain 92 As, β-chain 118 As	Stimulation of growth of ocytes and follicle
Thyrotropic hormone Hormon (TSH) polypeptide, α-chain, 92 aa, β-chain, 112 aa	Release of thyroxine (T_4 hormone) and of T_3 in thyroid gland
TSH-Releasing-Hormon peptide 3 aa	Formation in hypothalamus, stimulates synthesis and release of TSH in anterior pituiltary
Vasopressin peptide, 9aa	Formation in posterior pituitary, backresorption of water in the kidney, contraction of small blood vessels
Parythyroid hormone polypeptide, 84 aa	Formation in parathyroid gland, increase of Ca^{2+} in the blood, mobilization of Ca^{2+} from the bone

3.3.2
Hormone Analogs: Agonists and Antagonists

The modification of hormones can lead to compounds that are known as agonists or antagonists.

Antagonists are hormone derivatives that bind to a receptor but do not initiate signal transduction. Antagonists block the receptor and thus terminate signal transduction. Hormone antagonists find broad pharmaceutical and medical application since they specifically interfere with certain signal transduction pathways in the case of hormonal disregulation. Antagonists with a much higher affinity for a receptor than the unmodified hormone are medically very interesting. Such high-affinity antagonists require very low dosages in therapeutic applications. A few important antagonists and agonists of adrenaline are shown in Fig. 3.6. Propranolol is an example of a medically important hormone antagonist. Propranolol binds with an affinity three orders of magnitude greater than its physiological counterpart, adrenaline, on the β-adrenergic receptor. In this manner a very effective blockage of the adrenaline receptor is possible.

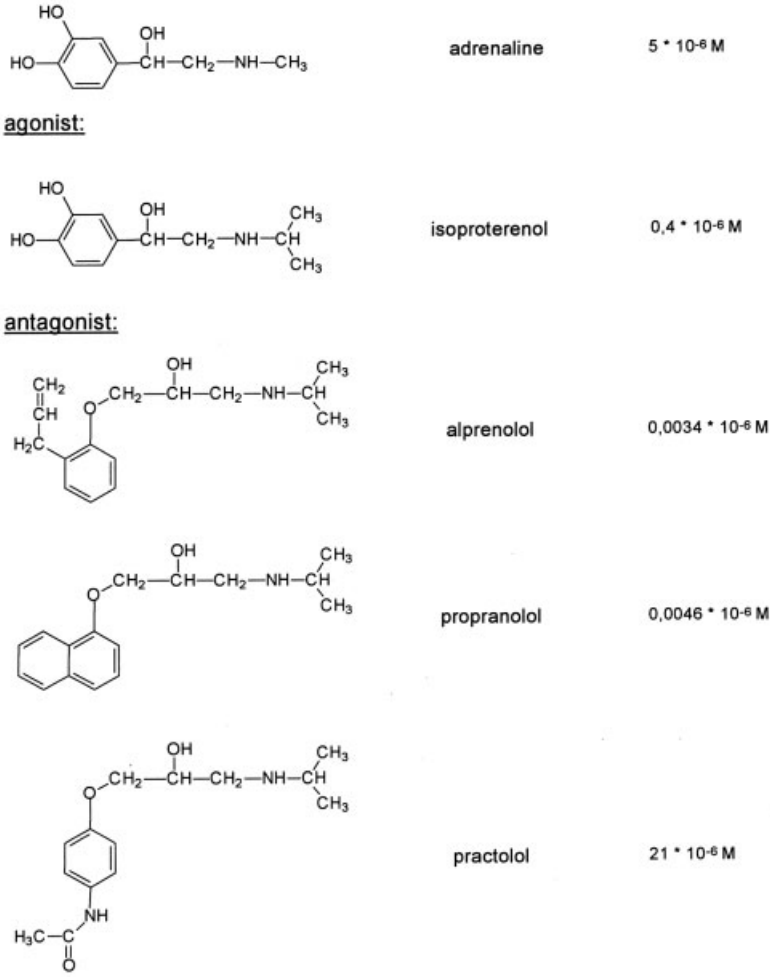


Fig. 3.6 Structure of important agonists and antagonists of adrenalin and their affinity for the β -adrenergic receptor (source: Lefkowitz et al., 1976).

Hormone analogs that bind specifically to a receptor and initiate the signal transduction pathway in the same manner as the genuine hormone are termed *agonists*. Application in research and medicine is found especially for those agonists which posses a higher affinity for a receptor than the underivatized hormone.

3.3.3
Endocrine, Paracrine and Autocrine Signaling

Various forms of intercellular communication can be discerned based on the range of the signal transmission (Fig. 3.7).

Endocrine Signaling
In endocrine signaling, the hormone is synthesized in specific signaling, or endocrine, cells and exported via exocytosis into the extracellular medium (e.g., blood or lymphatic fluid in animals). The hormone is then distributed throughout the entire body via the circulatory system so that remote regions of an organism can be reached.

Paracrine Signaling
Paracrine signal transduction occurs over medium range. The hormone reaches the target cells from the hormone-producing cell by passive diffusion. The producing cell must be found in the vicinity of the receiving cells for this type of communication. The signaling is rather local, and the participating signaling molecules are termed *tissue hormones* or *local mediators*. A special case of paracrine signal transduction is synaptic neurotransmission in which a nerve cell communicates with either another nerve cell or with a muscle cell.

Autocrine Signaling
In autocrine signaling, cells of the same type communicate with one another. The hormone produced by the signaling cell affects a cell of the same type by binding to receptors on these cells and initiating an intracellular signal cascade. If an autocrine hormone is secreted simultaneously by many cells then a strong response occurs in the cells. Autocrine mechanisms are of particular importance in the immune response (see Chapter 11).

3.3.4
Direct Modification of Protein by Signaling Molecules

A special case of signal transduction is represented by a class of small, reactive signaling molecules, such as NO (see Chapter 6.10). NO is synthesized in a cell in response to an external signal and is delivered to the extracellular fluid. Either by diffusion or in a protein-bound form, the NO reaches neighboring cells, and modification of target enzymes ensues, resulting in a change in the activity of these enzymes. NO is characterized as a mediator that lacks a receptor in the classical sense.

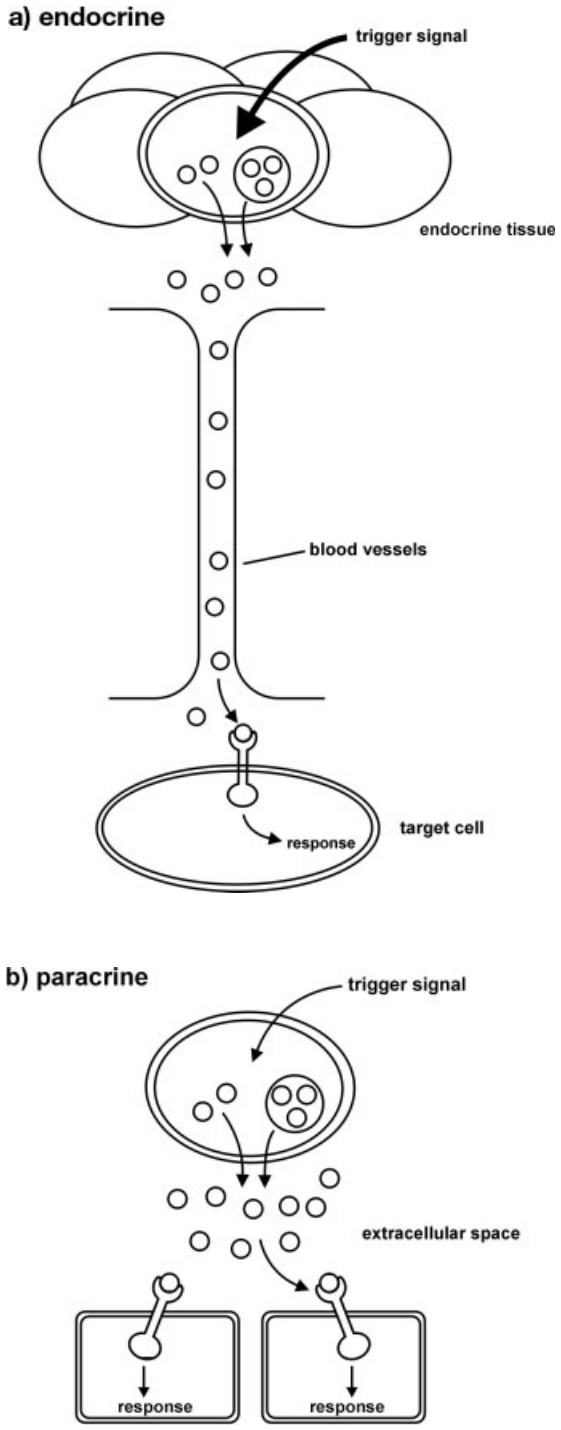


Fig. 3.7 Endocrinal, paracrinal and autocrinal signal transduction. a) endocrinal signal transduction: the hormone is formed in the specialized endocrinal tissue, released into the extracellular medium and transported via the circulatory system to the target cells. b) paracrinal signal transduction: the hormone reaches the target cell, which is found in close juxtaposition to the hormone producing cell, via diffusion. c) autocrinal signal transduction: the hormone acts on the same cell type as the one in which it is produced.

3.4
Hormone Receptors

3.4.1
Recognition of Hormones by Receptors

Hormones are usually produced by specialized cells and initiate a reaction in only a certain cell type. Only those cells that possess a cognate protein, the receptor of the hormone, can act as target cells. Receptors specifically recognize and bind the cognate hormone based on their chemical nature. The binding of the hormone to the receptor in the target cell induces an intracellular cascade of reactions at whose end lies a defined biochemical response. The pathway from receptor-bound signaling molecule to final biochemical response is complex and occurs under the participation of many proteins.

The receptors of the target cell can be divided into two classes: the membrane-bound receptors and the soluble cytoplasmic or nuclear localized receptors (Fig. 3.8).

Membrane-bound receptors are actually transmembrane proteins; they display an extracellular domain linked to an intracellular domain by a transmembrane domain. Binding of a hormone to the extracellular side induces a specific reaction on the cytosolic side, which then triggers further reactions in the target cell. The mechanism of signal transmission over the membrane will be discussed in more detail in Chapters 5, 8 and 11. Characteristic of signal transduction via membrane bound receptors is that the signaling molecule does not need to penetrate the target cell to activate the intracellular signal chain.

In the case of *intracellularly localized receptors* the hormone must enter the cell in order to be able to interact with the receptor. The hormone usually penetrates the target cell by passive diffusion. The nuclear receptors can be classified as ligand-controlled transcription activators. The hormone acts as the activating ligand; the activated receptor stimulates the transcriptional activity of genes which carry DNA elements specific for the receptor.

3.4.2
The Interaction between Hormone and Receptor

Receptors are the specific binding partners for signaling molecules; the former are able to recognize and specifically bind the latter based on their chemical structure. The binding and recognition are governed by the same principles and the same noncovalent interactions as those for the binding of a substrate to an enzyme, namely H-bonds, electrostatic interactions (including dipole-dipole interactions), van der Waals interactions and hydrophobic interactions. Signaling molecules bind their cognate receptors with an affinity greater than that usually observed for an enzyme and substrate.

The binding of a hormone to a receptor can in most cases be described by the simple reaction scheme:

$$[H] + [H] \rightleftharpoons [HR], \text{ with } K_D = \frac{[H] \cdot [R]}{[HR]}$$

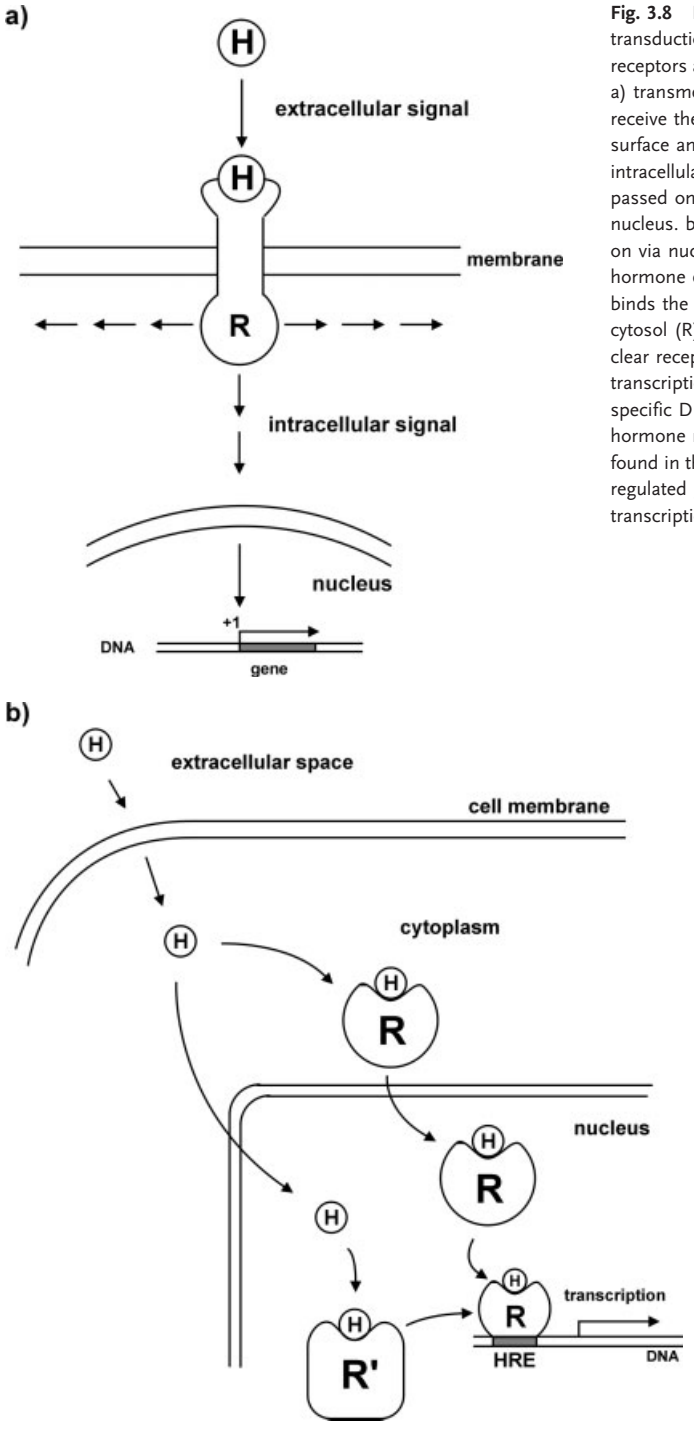


Fig. 3.8 Principles of signal transduction by transmembrane receptors and nuclear receptors. a) transmembrane receptors receive the signal on the cell surface and convert it into an intracellular signal that can be passed on until it reaches the nucleus. b) In signal transduction via nuclear receptors the hormone enters the cell and binds the receptor either in the cytosol (R) or nucleus (R'). Nuclear receptors act as nuclear transcription factors that bind specific DNA elements (HRE: hormone responsive element) found in the promotor region of regulated genes to control their transcription rate.

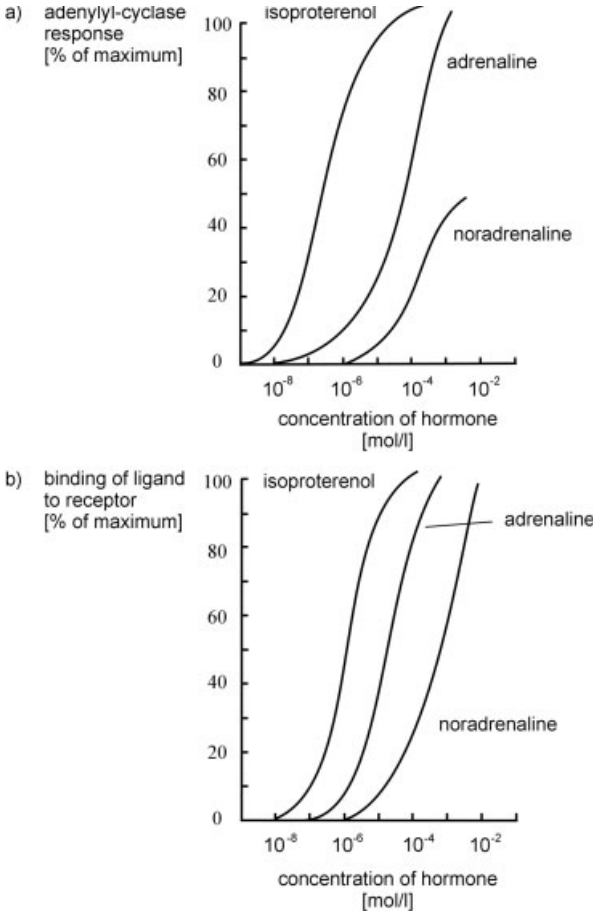


Fig. 3.9 Receptor binding and stimulation of the synthesis of cAMP by adrenaline, noradrenaline and isoproterenol. Adrenaline is the first member of a signal transduction chain that begins with the binding of of adrenaline to the β-adrenergic receptor. The signal is transmitted to adenylyl cyclase via a G-protein. The activated adenylyl cyclase forms cAMP, which serves as a diffusible intracellular messenger (see ch. 5& 6). The activation of cAMP synthesis and the binding to the β-adrenergic receptor as a function of hormone concentration in frog erythrocytes is shown in the figure above. a) A cell suspension is incubated with different concentrations of the three hormones, the cells are lysed and the adenylate cyclase activity determined. b) The degree of binding of adrenaline, noradrenaline and isoproterenol to the β-adrenergic receptor of frog erythrocytes as a function of hormone concentration. The curves show that the ability of a hormone to bind the β-receptor in a particular concentration range is highly correlated with the stimulation of adenylyl cyclase.The concentration of hormone required for half-maximal receptor binding is approximately the same as required for half-maximal adenylyl cyclase stimulation. Of the three ligands, isoproterenol binds the receptor the tightest and noradrenalin the weakest. Accordingly, stimulation of adenylyl cyclase is already observe at a very low isoproterenol concentration, while for the weaker binding noradrenaline much higher concentrations are required to stimulate the adenylyl cyclase. After Lefkowitz et al. (1976).

where [H] is the concentration of free hormone, [R] is the concentration of the free receptor, and [HR] is the concentration of hormone-receptor complex. The value for the equilibrium constant, K_D, usually lies in the range of 10^{-6} to 10^{-12} M. The binding curve for hormones and receptor are, according to the above simple equilibrium, hyberbolic in form.

Decisive for the intensity of the signal transmission is the concentration of the hormone-receptor complex, since the activation of the signal pathway requires this complex to be formed. The concentration of the hormone-receptor complex depends on the concentration of the available hormone, the affinity of the hormone for the receptor, and the concentration of the receptor. All three parameters represent, at least in principle, control points for signal transduction pathways. The variable signal, whose change is registered to thereby activate a signal transmission, is in most cases the concentration of the freely circulating hormone.

Synthesis and release of the hormone in endocrinic tissue is triggered in many hormonally regulated processes by an external signal. As a consequence, the concentration of the circulating hormone is increased, which implies an increased concentration of the hormone-receptor complex at the surface or within the target cell, and thus an increased activation of the downstream components of the signal pathway. For an efficient signal transduction, it is important that the change in concentration of the circulating hormone is approximately proportional to a change in the degree of occupation of the hormone receptor. This condition is fulfilled when the concentration of the hormone is in the range of the equilibrium constant, K_D.

The switch for the activation of an intracellular signaling pathway is in most cases an increase in the concentration of the freely circulating hormone. This leads to an increase in the concentration of the hormone-receptor complex, which results in an increased activation of subsequent reactions in the cell. The *concentration of the circulating hormone* is thus the main regulatory parameter in cellular communication. The relation between hormone concentration, binding of the hormone to the receptor, and subsequent reaction in the cell is illustrated in Fig. 3.9 for the case of adrenaline and the β-adrenergic receptor.

An increase in the hormone concentration is the main regulating element when a rapid activation of a signaling pathway is required. Hormones can be stored in the signal-producing cell in specialized organelles from which they can be quickly released when the appropriate external stimulus is received.

A modulation and regulation of signaling is also possible at the level of the receptor by varying the affinity of the receptor for the hormone or by specifically altering the concentration of the receptor. A change in the affinity of the receptor can be achieved by, for instance, phosphorylation of the receptor protein. The concentration of receptor available on the cell surface can be decreased either by degradation of the receptor or by internalization of the receptor (see Chapter 5). Both processes affect the intensity of the signal transduction on a long time scale.

3.5 Signal Amplification

Signal pathways commonly amplify the initial signal received by the receptor during the course of the signal transduction (Fig. 3.10). In many cases only a few molecules of a hormone are sufficient to initiate an enzymatic reaction in a cell, in which many substrate molecules are turned over. The extent of amplification, or amplification factor, varies greatly at the different levels of the signal transmission. An initial amplification often occurs at the level of the hormone-receptor complex. An activated receptor is capable of activating many downstream effector proteins.

The signal amplification at the level of the hormone-receptor complex depends upon many factors:

- **Life span of the hormone-receptor complex**
  The life span of the hormone-receptor complex is controlled primarily by the dissociation rate of the bound hormone.
- **Frequency of the reaction with the effector protein**
  An activated receptor can only transmit the signal further if it encounters an effector molecule. The frequency with which this occurs depends on the concentration and rate of diffusion of both components.
- **Deactivation of the hormone-receptor complex**
  The signal transmission by the hormone-receptor complex can be actively inhibited via covalent modifications (e.g., protein phosphorylation) which deactivate the hormone-receptor complex. Another mechanism for termination of signaling pathways is the internalization of the hormone-receptor complex. During internalization a section of the membrane, together with the proteins bound to it, is pinched off and transported into the interior of the cell. There the receptor can be returned to the cell membrane or be degraded. The internalization can affect the free receptor as well as the hormone-receptor complex.
- **Amplification of signaling during the visual process**
  One of the few examples for which the amplification factor at the level of the activated receptor could be determined is for the visual signal transduction pathway (review: Lamb, 1996). In the visual process a light signal is received by the photoreceptors, rhodopsin, of the rod cells which are then converted to the activated state, R*. The activated rhodopsin passes the signal on to the cognate G-protein, transducin, which in turn activates the next effector molecule, a cGMP phosphodiesterase. The phophodiesterase hydrolyzes cGMP to GMP. In the first step of the signal transduction, from the activated photoreceptor R* to transducin, there is a high amplification factor: *one activated rhodopsin molecule can activate 1000–2000 molecules of transducin per second*. No amplification occurs for the signal transmission from transducin to cGMP phosphodiesterase, since each activated transducin molecule only activates one phophodiesterase molecule. A further increase of the signal is found at the level of the activated phosphodiesterase, which rapidly hydrolyzes cGMP to GMP (k_{cat} = ca. 4000 s^{-1}).

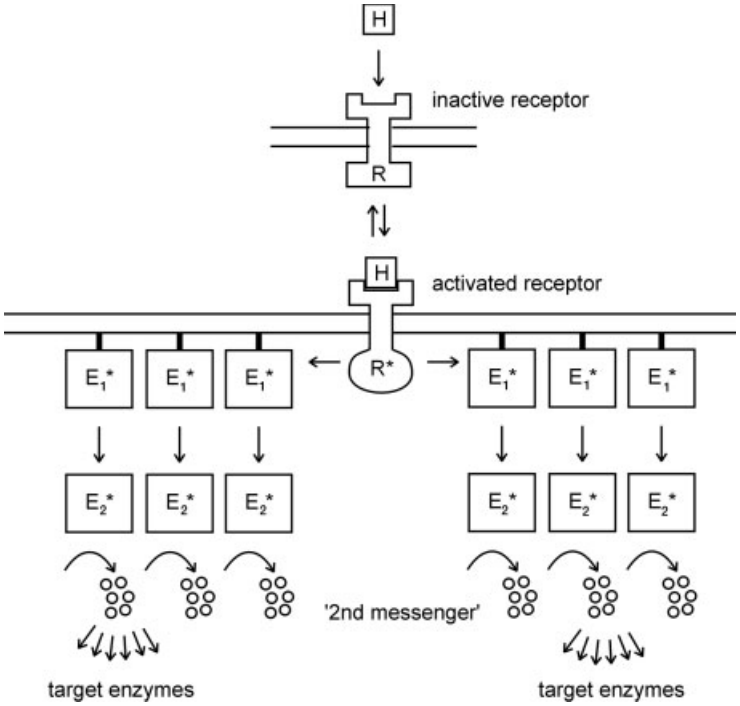


Fig. 3.10 Mechanisms of amplification in signal transduction pathways. Starting from an activated receptor R* many downstream effector molecules can be activated. E1* passes the signal on to E2, which is then converted to E2*. If E2 possesses enzymatic activity, and forms a diffusible messenger in its activated form, then a further amplification can occur at this step. The figure also emphasizes that, apart from the steps with a high amplfication factor, there are also steps that are not accompanied by amplification (here E1* → E2*).

The *lifetime* of the activated state of a signaling protein is an important regulatory point in the signal cascade. As will be discussed later in more detail for the regulatory GTPases (see Chapters 5 and 9), an extension or shortening of the activated state of a signaling protein can lead to an enhancement or attenuation of the signal transmission.

An exact determination and analysis of the amplification factor of signaling cascades *in vivo* is rarely possible. To determine the amplification factor, the life span of the activated state of a signaling protein, the concentration of the signaling protein and its cognate effector molecule, as well as the extent of deactivating processes in the cell must be known. These parameters are very difficult to determine experimentally. The concentration of the participating proteins is an especially elusive parameter. A further complicating factor is the membrane association of many signaling pathways, which hinders an accurate concentration determination.

3.6
Regulation of Inter- and Intracellular Signaling

The result of communication between the signaling and receiving cells is a defined biochemical reaction in the target cell. The nature and extent of this reaction depends on many individual processes that participate either directly or indirectly in signal transduction.

Beginning with the hormone-producing cell, the following processes are all contributing factors for hormonal signal transduction in higher organisms (Fig. 3.11):

1. biosynthesis of the hormone
2. storage and secretion of the hormone

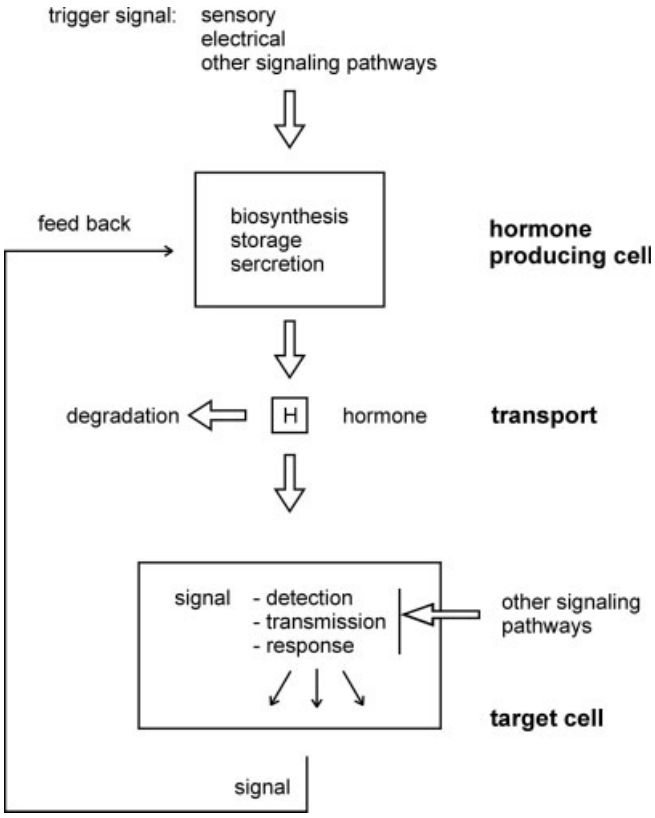


Fig. 3.11 Schematic representation of processes that may influence hormonal action in a cell. To note is the possibility for feedback in the framework of intercellular communication. A signal released in the target cell can regulate the hormone producing cell by, for example inhibiting the synthesis or secretion of the hormone. Furthermore, the possibility of a hierarchical structure and the mutual influence of different signaling pathways should also be noted.

3. transport of the hormone to the target cell
4. reception of the signal by the hormone receptor
5. transmission and amplification of the signal; biochemical reaction in the target cell
6. degradation and excretion of the hormone.

All of the above steps are subject to regulation. Thus, the effective concentration of a hormone at the target cell can be adjusted over a wide range according to numerous mechanisms. The biosynthesis of a hormone can, for example, be controlled by other signal transduction pathways. There are signals to trigger the secretion of stored hormones. The distribution of a hormone in an organism contributes to the accessibility of that hormone at a particular location. In addition, degradation and secretion of the hormone also plays an essential role in the effective concentration of the hormone in the cell.

The amount, activity and specificity of receptors at the target cell influences the extent of the final biochemical reaction. The induced reaction cascade in the target cell can be modulated at many positions by, for example, phosphorylation and subsequent changes in the activity of central proteins.

A signal transduction chain cannot be viewed as an isolated event within an organism, but should rather be interpreted in the context of other signaling pathways. The cell possesses a large repertoire of mechanisms by which the extent of signal transduction can be regulated and by which different signaling pathways communicate. Many of these mechanisms will be dealt with in detail in later chapters.

Every individual cell of a multicellular organism is programmed to react to the many external signals in a characteristic and specific manner. The reaction pattern of a cell type depends on the unique pattern of receptors and the corresponding coupled reaction pathways. This furthermore determines the function and morphology of the cell to a large extent. This pattern of regulation and networking of pathways is not constant during the course of development of an organism, but rather is subject to a genetically determined variation.

3.7
Membrane Anchoring and Signal Transduction

Transmission of signals over the cell membrane requires cooperation of the signaling proteins, each of which either exists as a transmembrane protein or is associated with the membrane. Extracellular signals are initially transmitted across the membrane with the aid of transmembrane proteins. In many cases the further signal transduction is tightly coupled to the membrane. This signal transduction, localized at the cytoplasmic side of the membrane, occurs with proteins whose function is tightly coupled with an association with the cell membrane.

An example of such a signal transduction can be found in the pathway leading to the production of cAMP, in which the signal is transmitted from a G-protein-coupled receptor to the cytoplasmic side of the membrane (see Chapter 5), where membrane-anchored heterotrimeric G-proteins become activated and transmit the signal

to downstream transmembrane adenylyl cyclases. The entire activation process occurs in tight association with the cell membrane. A tight coupling between the membrane and signal transduction process is also observed for signaling-linked enzyme activation when the substrate is localized in the membrane. Examples of such processes are found in phospholipases and PI3-kinases (see Chapter 6).

Signaling proteins are targeted to the inner face of the cell membrane mainly by the following ways:

- anchoring in the phospholipid bilayer by means of lipid anchors (see below)
- association with the membrane via binding of second messengers that are hydrophobic, e.g., phosphatidyl-inositol-trisphosphate (see Chapter 6)
- specific interaction with transmembrane proteins via protein modules like SH2 or PTB domains
- specific binding to membrane-targeting proteins that are stably associated with the membrane.

To achieve a stable and regulable membrane association of signaling proteins, a combination of these mechanisms may also be used.

A central and widely used tool for membrane anchoring is the post-translational attachment of hydrophobic residues, such as fatty acids, isoprenoids (see Fig. 3.12) or complex glycolipids (see Fig. 31.15) to specific amino acid side chains of target proteins. These lipid moieties of lipidated proteins favor membrane association by increasing the affinity of the protein to the membrane. Because of their hydrophobic nature, the membrane anchors insert into the phospholipid bilayer and thus mediate

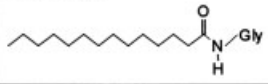
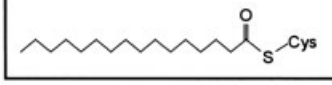
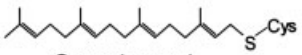
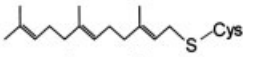
lipid	examples of modified proteins	site of modification
N-myristoyl 	heterotrimeric G-proteins (α-subunit), see chapter 5 cytoplasmic tyrosine kinases, see chapter 8	N-terminus
S-palmitoyl (S-Acyl) 	heterotrimeric G-proteins (α-subunit), G-protein-coupled receptors, see chapter 5 Ras proteins, see chapter 9	internal, no distinct consensus sequence
S-prenyl  Geranylgeranyl  Farnesyl	heterotrimeric G-proteins (γ-subunit) Ras proteins rhodopsin kinase see chapter 5	C-terminus

Fig. 3.12 Structure of lipid anchors and representative examples for lipid-modified signal proteins.

membrane association of the protein. To achieve a strong and stable membrane association more than one lipid anchor may be used.

The lipid anchors serve several functions in cell signaling. The main function is to promote membrane association of signaling proteins. Lipid anchors target proteins to the membrane, as is the case for the cytoplasmic protein tyrosine kinases, so that they can participate in membrane-associated signaling pathways. Importantly, the lipid anchors can be used in a dynamic way to recruit signal proteins in a regulated manner into signaling pathways. Furthermore, protein lipidation can mediate protein-protein association and/or stabilize protein conformations (review: Resh, 1999).

3.7.1 Myristoylation

Myristoylated proteins contain a saturated acyl group of fourteen carbons, myristoic acid (*n*-tetradecanoic acid) added cotranslationally via an amide bond to the amino group of the NH_2-terminal glycine residue. The consensus sequence for N-myristoylation is Gly-X-X-X-Ser/Thr (X: any amino acid), where the residue following the glycine is often a cysteine (review: Resh, 1999). A clumping of basic amino acids at the N-terminus can serve as an additional signal for myristoylation (Fig. 3.13). An example of this is c-Src (see Chapter 8), which displays a net charge of +5 at the N-terminal region. Clusters of basic residues at the C-terminus also aid to anchor the protein to the interface of the membrane, as observed in the example of the K-Ras protein (see Chapter 9 and Fig. 3.13). The basic amino acids help to anchor the protein to

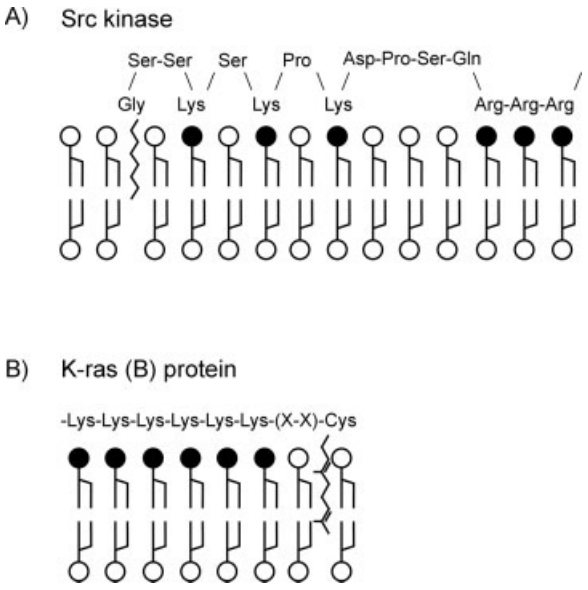


Fig. 3.13 Lipid anchors and basic regions as elements of the membrane association of proteins. Examples for proteins which exhibit basic residues near a lipid anchor. a) Src kinase (see Chapter 8) possess a myristoyl anchor at the N-terminus as well as a stretch of basic residues. b) In Ki-Ras proteins (see Chapter 9) there is a farnesyl residue at the C-terminus that serves as a lipid anchor, as well as a stretch of Lys residues. Negatively charged head groups of phospholipids are shown as filled circles. X: any amino acid.

the cell membrane via electrostatic interactions with the negatively charged head groups of the phospholipids.

Myristoylation is generally considered a constitutive process and a permanent modification. As shown below, the myristoic anchor may function as a switch during regulated membrane anchoring. Examples of myristoylated proteins are the cytoplasmic protein tyrosine kinases (family of the Src-kinases, Chapter 8), as well as the α-subunit of the heterotrimeric G-proteins (Chapter 5).

3.7.2 Palmitoylation

Palmitoylated proteins contain a long-chain fatty acid, such as palmitoic acid (*n*-hexadecanoic acid) connected to the protein via a labile thioester bond to cysteine residues. Other long-chain fatty acids like stearate and oleate have also been found to be incorporated in S-acetylated proteins. The thioester bond of S-acetylated proteins is less stable than the amide bonds of the myristate anchor. The lability conveys a reversible character to the modification and thus permits regulation of the membrane anchor (review: Dunphy and Linder, 1998). The reversible S-acetylation of signal proteins is thus a potential instrument for the modulation or regulation of signaling pathways. In the activation of G-protein-mediated signal pathways (see Chapter 5), the palmitoic acid anchor on the G_α-subunit is exchanged (review: Mumby, 1997). The distribution of the signal protein between the membrane and cytosol is possibly regulated via a cyclic acylation and deacylation. The enzymes involved in this process, the palmitoyl acyl transferases and the palmitoyl protein thioesterases, have been only partially characterized. According to the position of the acceptor cysteine, palmitoylated proteins can be divided into at least four classes (review: Resh, 1999). One class is dually acylated and carries both a myristoyl and a palmitoyl anchor. This dual acylation is encoded by the consensus sequence Met-Gly-Cys at the N-terminus.

Palmitoylation is, after myristoylation, the most common modification of the α-subunit of the heterotrimeric G-proteins (see Chapter 5). The use of both a myristate and a palmitate signal in Gα proteins is explained by the strong membrane binding mediated by two lipid anchors. Clearly, two fatty acids provide a stronger association with the membrane than only one could do. This explanation also pertains to dually palmitoylated proteins and to proteins that carry both a palmitate and a farnesyl anchor. The lipidation of cytoplasmic protein tyrosine kinases also includes both myristoylation and palmitoylation. H-Ras protein requires, apart from C-terminal farnesylation (see below), a palmitoyl modification in order to bind to the plasma membrane. In all the examples mentioned, the fatty acid anchors play an essential role in the signal transduction. A dynamic model has been proposed that accounts for the specific association of dually lipidated proteins with the membrane. According to the "kinetic bilayer trapping" hypothesis, proteins with a single lipid anchor only transiently and weakly associate with the membrane at many sites. Singly acylated protein that reaches the membrane will be rapidly palmitoylated by a membrane-associated palmitoyl transferase and will then remain stably attached to the membrane.

3.7.3
Farnesylation and Geranylation

Proteins with an isoprenoid modification possess either a C15-farnesyl residue or a C20-geranyl-geranyl residue. Both residues are bound via a thioester linkage to a cysteine residue. As with myristoylation, these are constitutive, stable modification performed by farnesyl or geranyl transferases.

The isoprenylation occurs at the Cys-residue of the consensus sequence Cys-A-A-X-COOH, whereby the nature of the C-terminal X-residue determines whether farnesylation or geranylation occurs (Fig.3.14). After the isoprenoid residue is appended, the three C-terminal residues are removed by a prenylation-dependent endopeptidase and the new COOH-group of the Cys-residue is methylated to increase the hydrophobicity of the C-terminus.

The isoprenoid modification can be found, among others, on the Ras protein and other members of the Ras superfamily (see Chapter 9), as well as with the α-subunit of

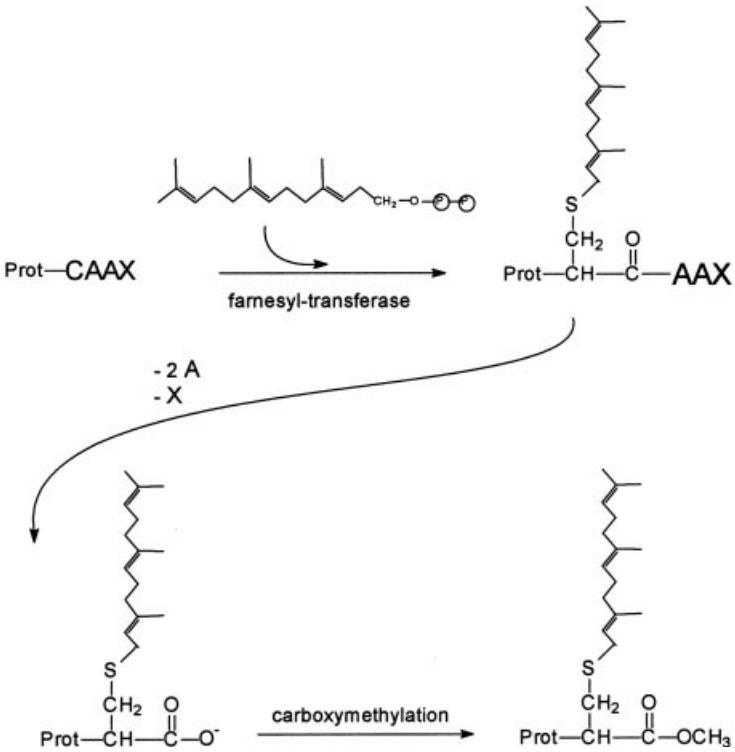


Fig. 3.14 Farnesylation at the C-terminus. The signal sequence for farnesylation is the C-terminal sequence CAAX. In the first step a farnesyl moiety is transferred to the cystein in the CAAX sequence. The farnesyl donor is farnesyl pyrophosphate and the responsible enzyme is farnesyl transferase. Subsequently, the three C-terminal amino acids are cleaved (A: alanine, X: any amino acid) and the carboxyl group of the N-terminal Cys-residue becomes methylated.

G-protein (see Chapter 5). The $\beta\gamma$-complex of G-proteins is also associated with the membrane via geranylation. A twofold geranylation is found on two Cys residues of the Rab protein (see Chapter 9).

In addition to promoting membrane association, other functional aspects of protein prenylation have been appreciated (review: Sinensky, 2000). Prenylation can also serve to mediate protein-protein interactions and has a role in protein trafficking.

3.7.4
The Glycosyl-Phosphatidyl-Inositol Anchor (GPI Anchor)

A lipid anchor can also be used to attach proteins to the extracellular side of cells. Cell surface proteins that mediate cell-cell interactions can be specifically anchored to the outward side of cells by means of a glycosl-phosphatidyl-inositol moiety.

The GPI anchor consists of a phospholipid with an appended glycosyl and ethanolamine residue in a complicated arrangement (Fig. 3.15). It is the most commonly employed anchor for the surface proteins of *Trypanosoma brucei* and prevalent in yeast. The function of the GPI anchor in mammals is not yet understood in detail.

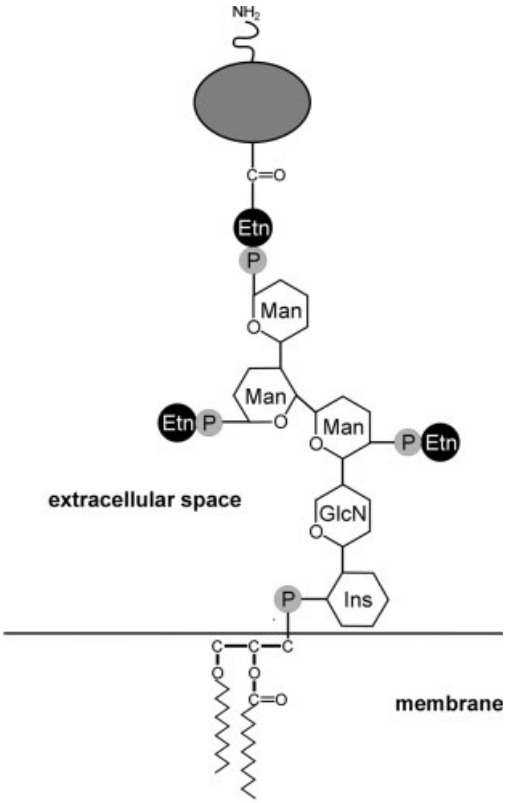


Fig. 3.15 Typical strucutre of a glycosyl phosphatidyl inositol (GPI) anchor. Ins: inositol; GlcN: 2'-amino, 2'-deoxy-glucose; Man: mannose; Etn: Ethanolamine, P: phosphate.

Cell surface proteins associated with the membrane via a GPI anchor participate in the uptake of nutrients, cell adhesion, and cell-cell interactions in the immune system. In T-lymphocytes, GPI-anchored proteins participate in signal transduction processes which lead to the activation of T-lymphocytes.

3.7.5 The Switch Function of Lipid Anchors

Lipid anchors can participate in a dynamic way in membrane anchoring and may thereby actively participate in cell signaling. We know of two ways by which the myristoyl anchor can function as a switch in cell signaling:

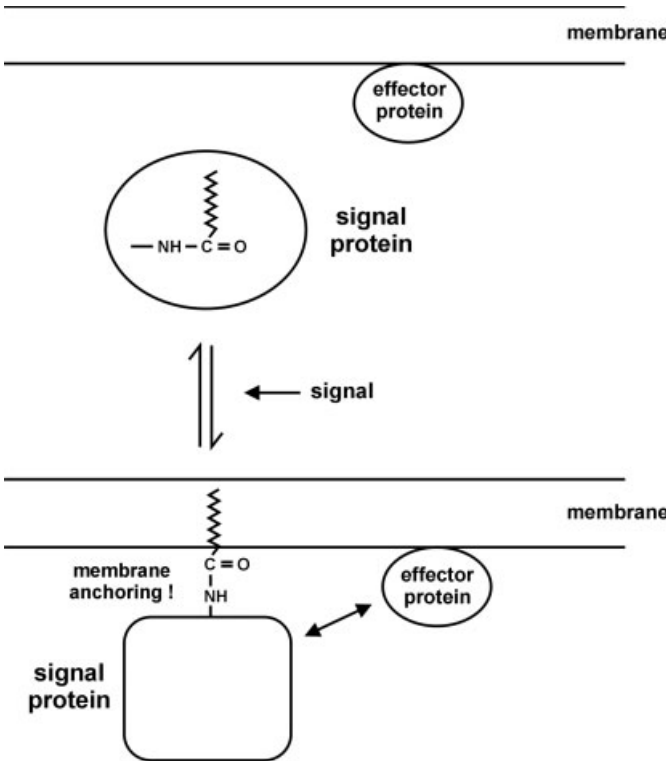


Fig. 3.16 Model of the switch function of the myristoyl anchor in signal proteins The myristoyl anchor of a signal protein can exist in a state accessible for membrane insertion or in a state buried in the interior of the protein. The transition between the two states may be controlled by specific cellular signals (e.g. Ca^{2+}, GDP/GTP exchange). In the membrane-associated form, interactions with membrane-bound effector proteins become possible and the signal can be transduced further.

- **Myristoyl-ligand switches**
  The orientation of the myristoyl moiety relative to the protein to which it is attached may be controlled by ligand binding. Depending on the presence of the ligand, the signaling protein can exist in a conformation where the lipid anchor is buried in the hydrohobic interior of the protein or in a conformation where it is exposed on the protein surface and accessible for membrane insertion. Membrane insertion of these proteins is reversible and is regulated by a specific ligand in a signal path controlled manner (Fig. 3.16). Examples are the Ca^{2+}-myristoyl switch of recoverin (see Section 6.7.3; Ames et al., 1997) and the GTP-myristoyl switch of the ARF-GTPase (Goldberg, 1998). In both cases ligand-induced conformational changes of the signaling protein are coupled to membrane binding.
- **Myristoyl-electrostatic switches**
  Another type of myristoyl switch has been reported for the MRACKS proteins which are substrates of protein kinase C (see Section 7.4). The membrane binding of the MARCKS proteins is mediated by myristate plus basic motif. Protein kinase C phosphorylation within the basic motif introduces negative charges into the positively charged region. This reduces the electrostatic interactions with the acidic phospholipids and results in displacement of the MARCKS proteins from the membrane and into the cytosol.

Reference

1 Ames, J.B., Ishima, R., Tanaka, T., Gordon, J.I., Stryer, L. and Ikura, M. (1997) Molecular mechanics of calcium-myristoyl switches. *Nature*, **389**, 198–202.

2 Dunphy, J.T. and Linder, M.E. (1998) Signalling functions of protein palmitoylation. *Biochim. Biophys. Acta*, **1436**, 245–261.

3 Goldberg, J. (1998) Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell*, **95**, 237–248.

4 Lamb, T.D. (1996) 'Gain and kinetics of activation in the G-protein cascade of phototransduction' *Proc. Natl. Acad. Sci. U.S.A.* **93**, 566–570

5 Lefkowitz, R.J.; Limbird, L.E.; Mukherjee, J. and Caron, M.G. (1976) 'The ß-adrenergic receptor and adenylate cyclase' *Biochem. Biophys. Acta* **457**, 1–55

6 Mumby, S.M. (1997) Reversible palmitoylation of signaling proteins. *Curr. Opin. Cell Biol.*, **9**, 148–154.

7 Resh, M.D. (1999) Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim.Biophys. Acta*, **1451**, 1–16.

8 Sinensky, M. (2000) Functional aspects of polyisoprenoid protein substituents: roles in protein-protein interaction and trafficking. *Biochim. Biophys. Acta*, **1529**, 203–209.

4
Signaling by Nuclear Receptors

Nuclear receptors regulate gene expression in response to binding small lipophilic molecules and are thereby involved in the control of a diversity of cellular processes. These proteins are ligand-activated transcription factors that are localized in the cytoplasm and/or in the nucleus. The ligands pass the cell membrane by simple diffusion and bind to the cognate receptors in the cytoplasm or in the nucleus. By binding to DNA elements in the control regions of target genes the ligand-bound receptor influences the transcription of these genes and thus transmits hormonal signals into a change of gene expression.

4.1
Ligands of Nuclear Receptors

The naturally occurring ligands of nuclear receptors are lipophilic hormones, among which the steroid hormones, the thyroid hormone T3, and derivatives of vitamin A and D have long been known as central regulators. These hormones play a significant role in metabolic regulation, organ function, and development and differentiation processes. Following formation and secretion in specific tissues, the endocrine organs, the hormones are distributed in the organism via the circulation and enter cells passively by diffusion. In recent years it has been recognized that intracellularly formed lipophilic metabolites can also serve as ligands for nuclear receptors and can regulate gene expression through their binding to nuclear receptors. These compounds include prostaglandins, leukotrienes, fatty acids, cholesterol derivatives, bile acids, and even benzoates. The most important natural ligands of the nuclear receptors are shown in Fig. 4.1; the cognate receptors and their DNA elements are summarized in Table 4.1 (see Mangelsdorf, 1995, and Aranda and Pascual, 2001).

Generally, the ligands of the nuclear receptors are small, rigid hydrophobic compounds with polar parts at one end, allowing binding to the receptor via Van der Waals and hydrophobic interactions and hydrogen bonds (see Section 4.3.4).

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

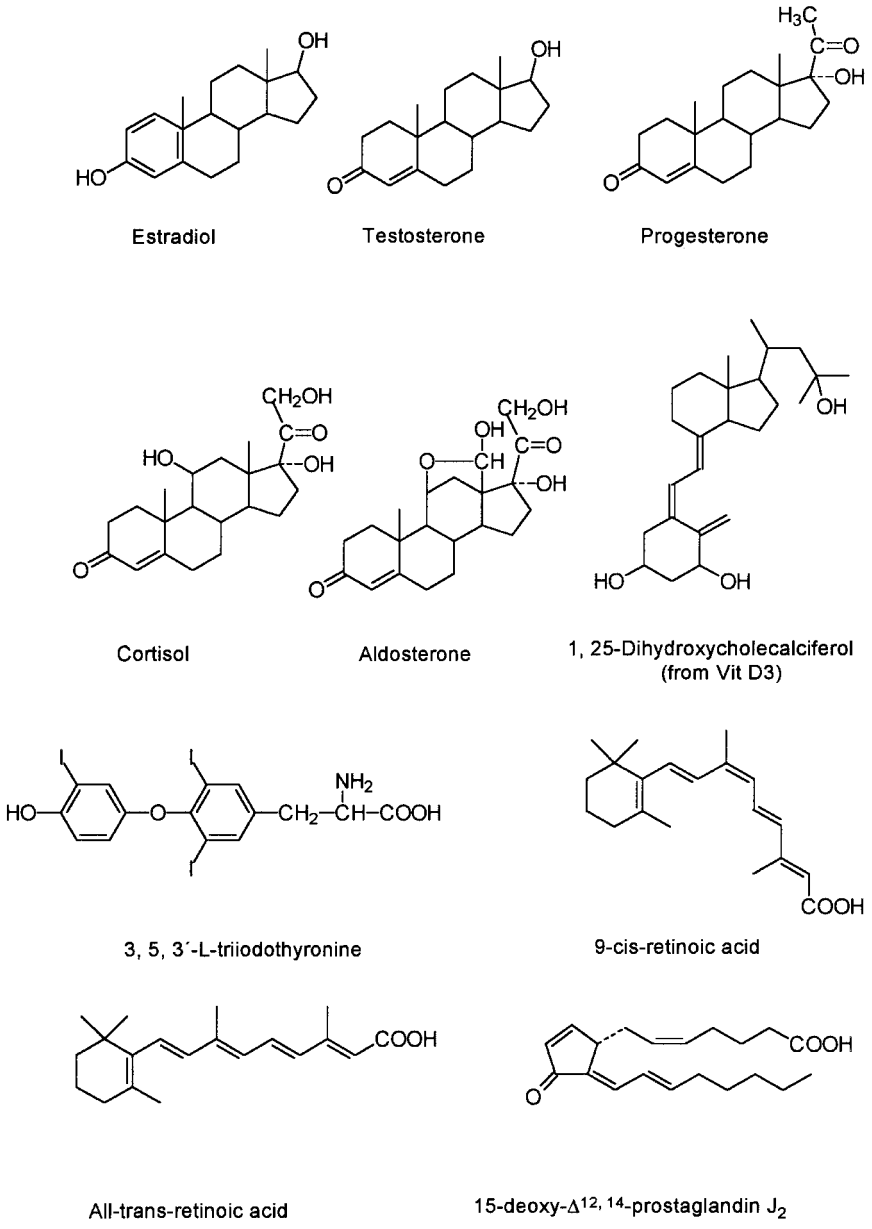


Fig. 4.1 Natural ligands of nuclear receptors.

Tab. 4.1 Ligands and structure of HREs of selected nuclear receptors from mammals.

Receptor	Hormone	Half site sequence	configuration of HRE
Glucocorticoid receptor, GR	Cortisol	AGAACA	IR-3
Mineralcorticoid receptor, MR	Aldosterone	AGAACA	IR-3
Progesterone receptor, PR	Progesterone	AGAACA	IR-3
Androgen receptor, AR	Testosteron	AGAACA	IR-3
Estrogen receptor, ER	Estrogen	RGGTCA	IR-3
Farnesoid receptor, FXR	Farnesoids	AGGTCA	IR-1, DR-5
Thyroid hormone receptor, T3R	3,5,3'-L-triiodothyronine, T_3 hormone	RGGTCA	IR-0, DR-4, ER-6,8
Receptor for vitamine D_3, VDR	1,25-Dihydroxy vitamine D_3	RGKTCA	DR-3
Receptor for 9-cis retinoic acid, RXR	9-cis retinoic acid Terpenoids	AGGTCA	DR-1
Receptor for all-trans retinoic acid, RAR	all-trans retinoic acid	AGTTCA	IR-0, DR-2,5; ER-8
Peroxisome proliferator-actived receptor, PPAR	15-Deoxy-Δ12, 14-prostaglandin J_2	AGGTCA	DR-1
COUP-TF (α, β, γ)	?	RGGTCA	DRs, IRs
NGFI-B (α, β, γ)	?	AAAGGTCA	NR, DR-5, IR-O
ROR (α, β, γ)	?	YYCYRGGTCA	NR

Abbreviations: IR: "inverted repeat"; DR: "direct repeat"; ER: "everted repeat"; NR: "no repeat". Numbers of HREs give the number of pairs separating the half-sites. R: purine. Y: pyrimidine. NGFI: Nerve growth factor induced receptor; COUP-TF: CHicken ovalbumin upstream promoter transcription factor; ROR: Retinoic acid related orphan receptor. α, β und γ are receptor subtypes coded by distinct genes. After Mangelsdorf et al., 1995.

4.2 Principles of Signaling by Nuclear Receptors

Signal transduction by nuclear receptors is shown schematically in Fig. 4.2. The nuclear receptors are localized in the cytosol and/or nucleus. Many of the natural ligands of nuclear receptors are lipophilic hormones that enter the cell in a passive manner or by active transport mechanisms. A transmittance of the signal at the cell membrane is not necessary, as is the case in the signaling via transmembrane receptors. Once inside the cell, the hormone ligand binds the cognate receptor which is localized in the cytosol and/or in the nucleus. The hormone binding activates the transcription regulation function of the receptor. In the case of the cytosolic receptors, the hormone binding induces translocation into the nucleus where the hormone-receptor complex binds a cognate DNA element termed *hormone-responsive element, HRE*, and alters the transcription of the target gene.

Apart from the classical, well-known hormones listed in Fig. 4.1, other compounds are also used as signaling ligands for the activation of certain nuclear receptors. These ligands may be synthesized intracellularly as normal metabolites such as fatty acids and bile acids and they may be derived from foreign lipophilic substances like drugs.

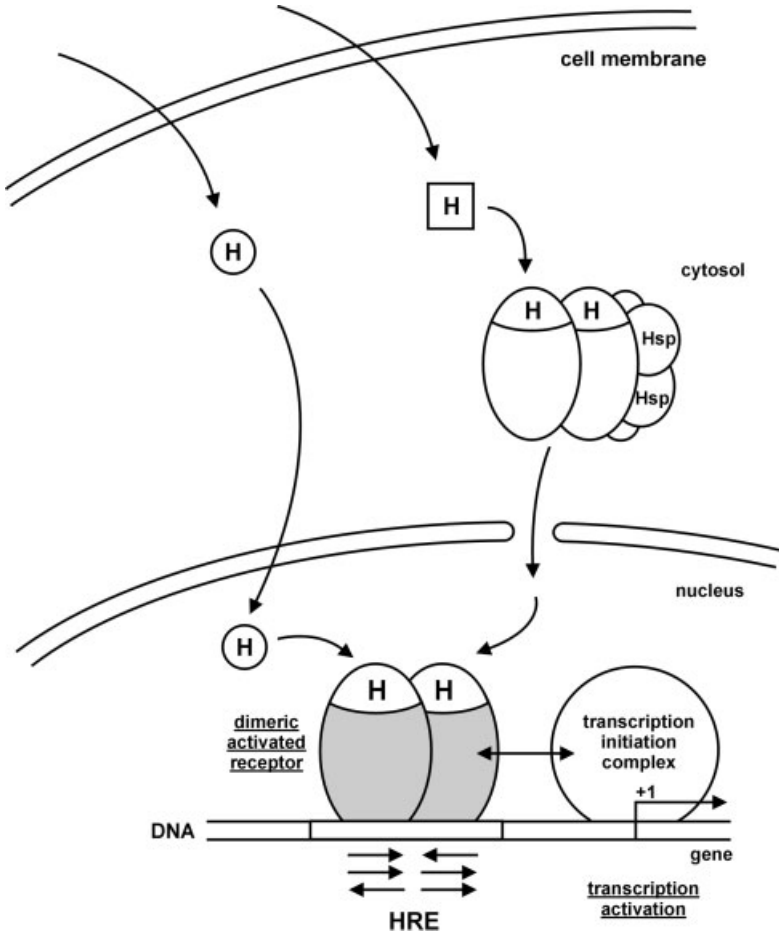


Fig. 4.2 The principle of signal transduction by nuclear receptors. Nuclear receptors are ligand-controlled transcritpion factors that bind cognate DNA sequences, or hormone responsive elements (HRE). The hormone acts as a regulating ligand. Most nuclear receptors bind their cognate HRE, which tend to be symmetrically organized, as homo- or heterodimers. The DNA-bound, activated receptor stimulates transcription initiation via direct or indirect protein-protein interactions with chromatin and he transcription initiation complex. The arrows demonstrate the different possible configurations of the HRE (see also Section 4.3.1). H: hormone; Hsp: heat shock protein.

The cognate receptors, e.g., the PPAR and the FXR, are quite promiscuous with respect to the nature of the ligand and have been shown to be able to bind a broad range of lipophilic ligands. This type of receptors is thought to be involved in the homöostasis of metabolism and in the detoxification of foreign substances.

Detailed reviews of the structure and function of nuclear receptors are found in Mangelsdorf et al. (1995), Perlman and Evans (1997), and Aranda and Pascual

(2001). The nuclear receptors are ligand-controlled transcriptional regulators that function as activators or repressors of transcription. They are DNA-binding proteins which bind specifically to cognate DNA elements, the HREs, located in the control regions of target genes. Depending on the receptor, DNA binding is linked to either *activation* or *repression* of the neighboring genes. Generally, ligand binding to the receptor activates transcription. Rare cases are, however, known where ligand binding induces a repression of transcription. Receptors that are localized in the cytosol are translocated upon ligand binding into the nucleus where they bind to their HREs and activate transcription. We also know of receptors that have a repressive effect on transcription when bound to DNA. In this case, ligand binding to the DNA-bound receptor relieves the repression, and transcription is activated.

In addition to ligand binding, other ways of activating nuclear receptors have been discovered. Phosphorylation by kinase signaling pathways as well as binding of other regulatory proteins like cyclin D1 (see Section 13.2.3) have been reported (Lin et al., 2002), indicating a linkage of other signaling pathways to nuclear receptor signaling.

In comparison to signaling pathways which utilize transmembrane receptors (see Chapters 5, 8, 11), signaling via nuclear receptors is of relatively simple structure. The pathways lead directly, with only a few participating protein components, from the extracellular space to the level of transcription in the nucleus. Up to the contact with the transcription machinery, most important protein components of the signal pathway are known and well characterized. However, because of the complexity of transcriptional regulation, we still do not fully understand how transcription activation by the nuclear receptors is achieved. This function has to be seen in the context of chromatin-bound DNA and its reorganization upon activation by transcriptional regulators. A multitude of proteins participate in the communication between transcriptional activators and the basal transcription machinery during transcription initiation. These proteins have an essential function in nuclear receptor signaling too. Coactivators, corepressors and mediators (see Section 1.4.4.2) have been recognized to be crucial players that transmit the signal from the activated nuclear receptor into a stimulation of transcription. As discussed in Section 1.4.4, the structural and functional basis of the function of these protein complexes is still incompletely understood. Part of their function is assumed to form a bridge to the basal transcription apparatus. The enzymatic activities of these proteins like histone acetylation and deacetylation are thought to play an essential role as well.

Regulation processes which act at the level of transcription are slower than those which affect the activity of already existing enzymes. The latter process is used mainly to induce a rapid reaction in an organism to an external signal or to achieve a rapid communication between cells. Signal transduction by nuclear receptors, on the other hand, is mostly intended to achieve long-term changes in the activity of enzymes, where the speed of the regulatory process is not very prominent. Regulation with the aid of nuclear receptors is therefore used mainly to adapt the activity of key metabolic enzymes to modified external conditions or a change in the function of an organism.

Signaling via nuclear receptors is paramount for the development and differentiation of higher organisms, since these processes do not require a rapid response but do

require long-lasting functional changes. Accordingly, many hormones of the signal transduction pathways involving nuclear receptors participate in the development and differentiation of organs. Examples are the sexual hormones, the thyroid hormone T_3, the D_3 hormone, and retinoic acid.

4.3 Classification and Structure of Nuclear Receptors

The first nuclear receptors to be characterized were those for the glucocorticoids, for estrogen, and for progesterone. Many other nuclear receptors could be identified based on common structural and functional features, and the nuclear receptors are now grouped into a large superfamily with at least six different subfamilies (Robinson-Rechavi et al., 2003). The nuclear receptors display a high degree of homology at the level of amino acid sequence, which indicates that they operate with similar functional principles. For some receptors the cognate hormone and their function in the cell remain unknown. Such *"orphan receptors"* are usually identified by sequence homology and with the help of screening techniques employing DNA probes based on known receptors. Possibly some of the orphan receptors do not have natural ligands and function in a ligand-independent manner. The most important representatives of the nuclear receptors are summarized in Table 4.1.

Like other transcriptional regulators, nuclear receptors exhibit a modular structure with different regions corresponding to autonomous functional domains. At the level of the primary structure the steroid hormone receptors can be divided into five different domains (Fig. 4.3), each with specific functions.

A typical nuclear receptor contains a variable N-terminal A/B region, a highly conserved domain responsible for the *DNA-binding* (region C), a linker region D containing *nuclear localization signals*, the *ligand-binding and dimerization* domain (region E), and an F region. Domains responsible for transactivation are found in the A/B region (activation function 1, AF-1) in the E region and in the F-region (activation function 2, AF-2).

Especially the domains for DNA binding and ligand binding display characteristics of independently folding structural units. The single domains can therefore be exchanged using recombinant DNA technology and their functionality in other protein environments examined in "domain-swapping" experiments (see Fig. 1.24).

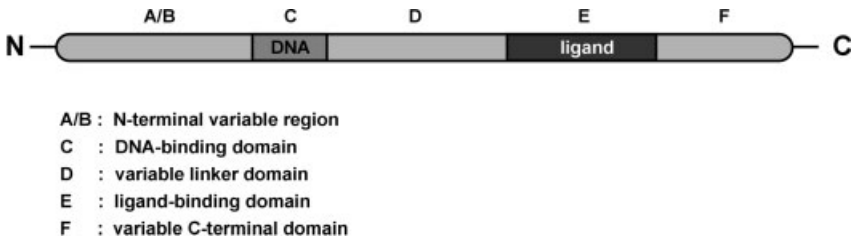


Fig. 4.3 Domain structure of the nuclear receptors. Functional domains of nucear receptors are portrayed in a one dimensional, linear fashion.

4.3.1
DNA-Binding Elements of Nuclear Receptors, HREs

The steroid hormone receptors are sequence-specific DNA-binding proteins whose cognate DNA elements are termed "hormone-responsive elements" (HREs). The HREs known to date possess a common structure. They are composed primarily of two copies of a hexamer sequence. In Table 4.1 are listed the hexamer sequences of the HREs of important nuclear receptors.

The identity of an HRE is determined by the sequence, polarity, and distance of the hexamers. Mutation and duplication of an ancestral recognition sequence have al-

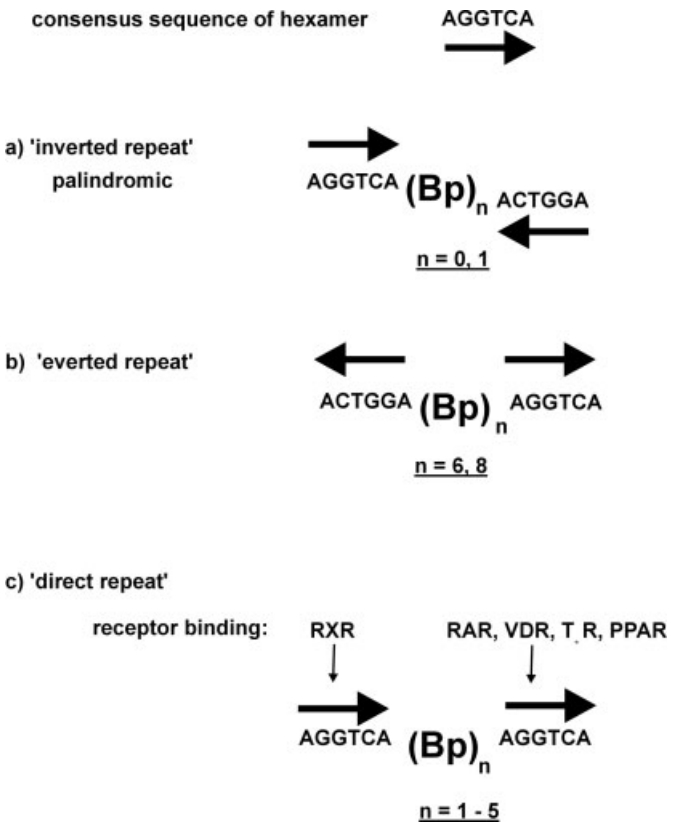


Fig. 4.4 HRE structure of the RXR heterodimer. Shown is the consensus sequence of the HREs of the RXR heterodimers (see Fig. 4.7) and the different possible arrangements of the hexameric half-site sequences. The hexamers can be arranged palindromically as inverted repeats (a), as everted repeats (b), or as direct repeats (c). n indicates the number of base pairs that lie between the two hexamers. RXR: receptor for 9-cis-retinoic acid; RAR: receptor for all-trans retinoic acid; T_3R: receptor for the T_3 hormone; PPAR peroxisome proliferator activated receptor; VDR: receptor for vitamin D_3.

lowed the creation of many and various DNA elements during the course of evolution, whose sequence, polarity and distance is characteristic for a given hormone receptor or receptor pair. The half-site of an HRE can be arranged as a palindrome, an inverted palindrome or a direct repeat. For a given receptor, optimal spacings of the half-sites exist, and the number of base pairs between the half-sites is another charactistic feature of a HRE. Figure 4.4 illustrates, on the example of the HREs for the RXR heterodimer (see Section 4.7), the various configurations of an HRE.

The receptors bind to the cognate HRE mainly as dimers, allowing the formation of *homodimers* as well as *heterodimers* between various receptor monomers. We know of very few nuclear receptors whose HRE contains only a single copy of the recognition sequence. These receptors bind as monomers to the cognate HRE.

Based on the subunit structure of DNA-bound receptors and on the structure of the HREs, four classes of nuclear receptors can be distinguished (Fig. 4.5).

Dimers of the Steroid Hormone Receptors

The HREs of the steroid hormone receptors possess a *palindromic structure*, comparable to the DNA-binding elements of procaryotic repressors (see Fig. 4.5a). The glucocorticoid receptor, for example, binds as a homodimer to the twofold symmetrical recognition sequence, whereby the receptor is already dimerized in solution. In complex with the DNA, each subunit of the dimer contacts one half-site of the HRE. As a consequence of the twofold repeat of the recognition sequence, a high affinity binding of the receptor dimer results (compare Section 1.2.4).

Heterodimers Containing RXR

The DNA-binding elements of the nuclear receptors for all-trans retinoic acid, for 9-cis retinoic acid, for the T_3 hormone and for the vitamin D_3 hormone usually exhibit a direct repeat of the recognition sequence, resulting in the formation of heterodimers on the DNA (Fig. 4.5b). One of the partners in the heterodimer is always the receptor for 9-cis retinoic acid, RXR, which usually occupies the 5' side of the HRE.

Of particular importance for receptor binding on HREs with direct repeats of the recognition sequence is the fact that the hexamers of these HREs are arranged head to tail and thus require a *polar arrangement* of the receptor dimers.

HREs of this type display a further unique characteristic: half-site sequences can be used to create different HREs by varying only the *spacing* between the repeats. The spacing can vary between 1 and 5 bp of any sequence, so that with one given repeat five different HREs can exist (see also Fig. 4.4). In this case, the spacing between the recognition sequences determines which hetero- or homodimer can form. Further multiplicity is achieved by combining different half-site sequences in a single HRE.

Dimeric Orphan Receptors

The "orphan receptors" derive their name from the fact that the cognate hormones for these receptors were originally unknown or little understood. Orphan receptors can bind as homodimers to recognition sequences arranged as direct repeats (Fig. 4.5c). The receptor for 9-*cis* vit A acid, which also binds as a homodimer an HRE with two half-sites in direct repeats, is also considered a member of this class of receptors.

Monomeric "Orphan Receptors"

Orphan receptors are also known (e.g., Retinoic Z receptor, RZR; NGF-induced clone B, NGFI-B) which bind as monomers to asymmetric recognition sequences (Fig. 4.5d). The function of these receptors is still poorly characterized.

4.3.2 The DNA-Binding Domain of Nuclear Receptors

The binding of receptors to their cognate HRE occurs via a DNA-binding domain, which are largely independent folding domains. Within the family of nuclear receptors, the DNA-binding domain is the most conserved structural element and is located in region C of the primary structure (see Fig. 4.6a). The DNA-binding domain pos-

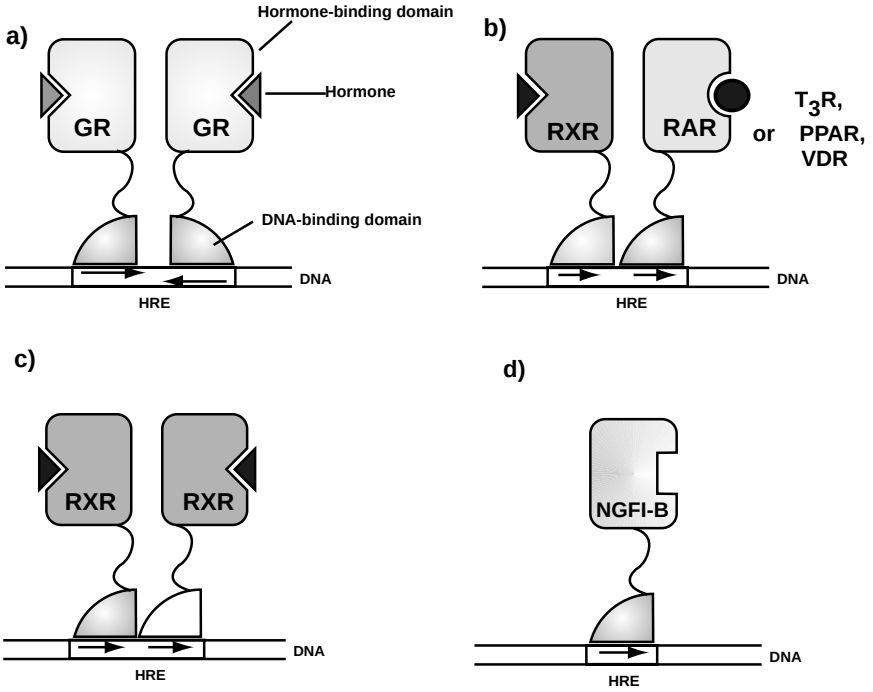


Fig. 4.5 Oligomeric structure of nuclear receptors and structure of the HREs. The nuclear receptors can be subdivided into four groups based on strucutres of the receptors and HREs. Shown above are some representative examples. a) binding of a homodimeric receptor to a two-fold symmetric palindromic DNA element, GR: gluccocorticoid receptor. b) binding of a heterodimeric receptor to a DNA element with direct repeats of the recognition sequence, whereby the 5' side of the HRE is occupied by the receptor for 9-cis retinoic acid (RXR). RAR: receptor for all-trans retinoic acid, T_3R: receptor for the T_3 hormone; PPAR: peroxisome proliferating activated receptor; VDR: receptor for vitamin D_3. c) binding of RXR as a homodimer to an HRE with direct repeat of the recognition sequence. d) binding of a monomeric receptor to an asymmetric recognition sequence, NGFI-B: nerve growth factor induced receptor, is involved in the regulation of enzymes of steroid biosynthesis. After Mangelsdorf et al., 1995.

sesses structural elements that mediate the specific recognition of the HRE, as well as for the dimerization of the receptor on the HRE. The core of the DNA-binding domain includes a span of 70–80 amino acids, in which all information for the specific recognition of the cognate half-site is contained.

In the core of the DNA-binding domain are two Zn_2Cys_4-motifs. As shown in Fig. 4.6, the Zn-motifs serve to position a recognition helix in the major groove of the DNA. It is not involved in any direct interactions with the bases of the DNA. Via the recognition helix specific contacts are formed with the hexamer half-site of

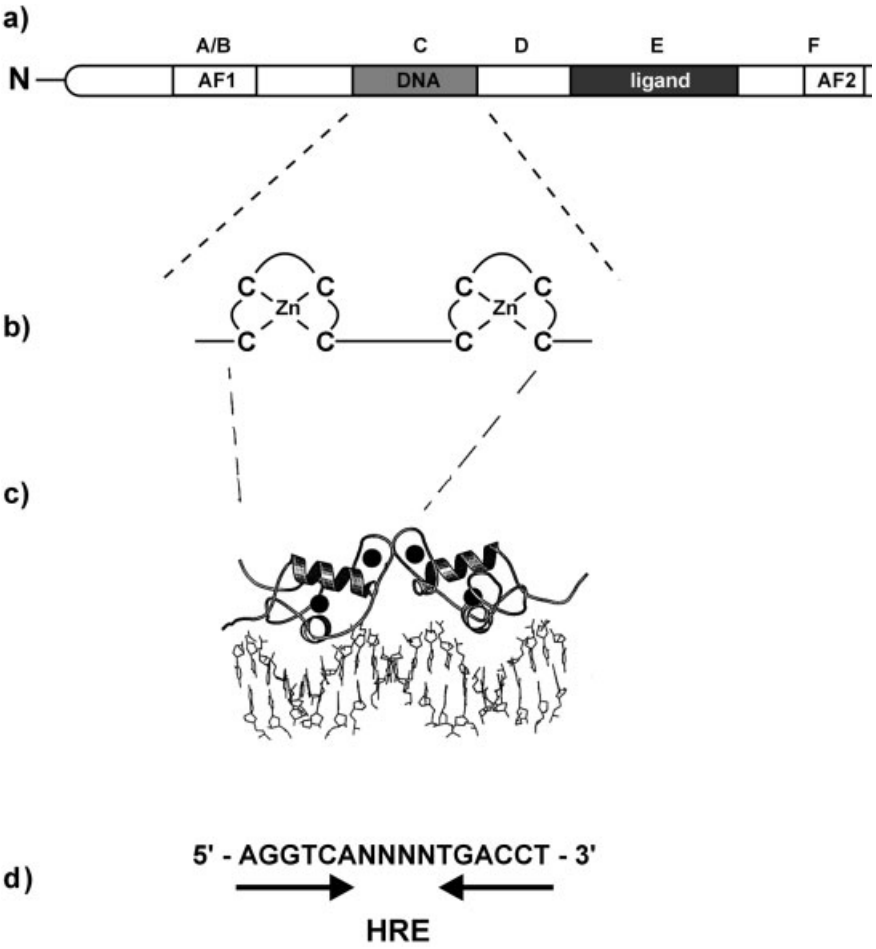


Fig. 4.6 Functional domains, DNA binding and HRE structure of nuclear receptors on the example of the glucocorticoid receptor, GR. a) domain structure of GR. AF1, AF2: domains that mediate the stimulation of the transcription. b) schematic representation of the two Zn^{2+}-Cys4 binding motifs of the DNA binding domains. c) Complex formation between the dimeric DNA binding domains of GR and the HRE. The black spheres represent Zn^{2+} ions. After Luisi et al., 1996. d) Consensus sequence and configuration of the HRE elements of GR.

the HRE. The Zn-motifs of the nuclear receptor are an integral part of the overall structure of the DNA-binding domain and do not represent an independent structural element, as is the case with the transcription factors TFIIIA or ZIF268 (compare Section 1.2.1 as well as Figs. 1.4 and 1.5). The two Zn-motifs assume nonequivalent positions in the DNA-binding domain. While the N-terminal Zn-motif participates in the positioning of the recognition helix and the interactions with the sugar-phosphate backbone, the C-terminal Zn-motif serves to impart a certain functionality to the dimerization surface and to contact the phosphate backbone of the DNA.

4.3.3
HRE Recognition and Structure of the HRE-Receptor Complex

It is a characteristic feature of the structure of HREs that the recognition sequences of different receptors are actually very similar, and that they differ primarily in the polarity and spacing of the sequences. It is therefore not surprising that the structure of the receptor dimers is accurately adapted to the exact orientation and spacing of the recognition sequences. High-resolution structures of DNA-bound homo- and heterodimers have confirmed this aspect of receptor-DNA recognition.

In the case of the palindromically arranged recognition sequences, the binding occurs mostly via pre-formed homodimers of the receptor in solution. Sequence and spacing of both recognition sequences in the HRE are highly complementary to the binding surface of the recognition helix, as well as to the spacing between the DNA-binding domains of the dimeric receptor. The spacing of the recognition sequences is firmly predetermined by the structural elements of the receptor participating in the dimerization. The geometry of the dimer determines, furthermore, what spacing of the recognition sequences is tolerated in the formation of a high affinity complex.

For direct repeat HREs, the spacing of the two half-sites is often the decisive, if not the only, element based on which the receptor (homodimer or heterodimer) recognizes its own HRE and discriminates against related HREs. The solution of the structures of heterodimer-DNA complexes (review: Steimetz et al., 2001) has shown how these receptors can distinguish between highly related HREs (Fig.4.7). As an example, the structure of a DNA-bound receptor-heterodimer composed of the DNA-binding domain of RXR and the T_3-receptor (Rastinejad et al., 1995) is given in Fig. 4.7.

The HRE sequence used for the structure determination ($AGGTCA(N)_4AGGTCA$) consists of two identical hexamers in direct repeat and separated by 4 bp (D-4 arrangement). The heterodimer RXR-T_3R binds in a polar manner on the HRE, with RXR occupying the 5' side of the HRE. Both hexameric sequences lie on the same side of the DNA double helix and are contacted by an α-helix of each of the receptors in a nearly identical manner. In the complex, the DNA-binding domain of T_3R interacts more with the half-site than does RXR. Different structural elements of each of the monomeric receptors are involved in the dimerization process, leading to the polar configuration of the monomers on the DNA.

The structure determination confirmed the importance of the spacing of the two hexamers as a discrimination factor in an impressive manner. A spacing of only 3

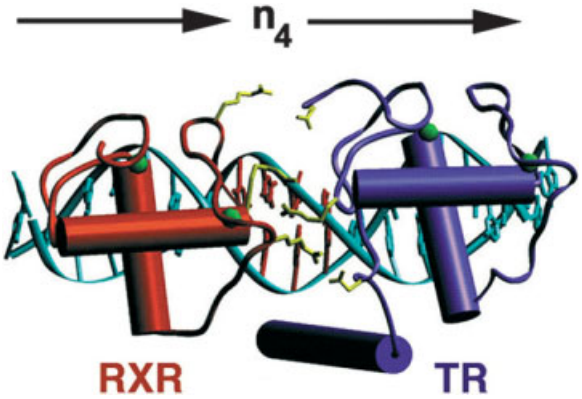


Fig. 4.7 Structure of the RXR-T_3R heterodimer in complex with DNA. Illustrated is a complex between the DNA binding domain of the RXR-T_3R heterodimer and an HRE with direct repeats of the sequence AGGTCA separated by 4 base pairs. The two receptor subunits contact the hexameric sequences with a recognition helix in a manner very similar to that of the gluccocorticoid receptor (see Fig. 4.6). The Zn atoms are drawn as green spheres. The figure illustrates the polarity of the binding of the two subunits. The interaction between the two subunits is mediated mainly via an extension of the C-terminal DNA binding domain (bottom) of the T_3R. A greater or smaller distance between the two hexamers of the HRE would act contrary to the interaction between the two subunits as shown. After Rastinejad, 2001.

nucleotides between the two hexamers would lead to steric overlap of both receptors; a high affinity, cooperative binding would not be possible. With a spacing of more than 4 nucleotides a high affinity complex could also not be formed because of the relative rigidity of the two monomers.

The dimeric structure of the receptors thus plays an important role in the discrimination between closely related hexamers. The structural elements of the receptors which participate in the dimerization ensure that the recognition helices assume a defined mutual spatial arrangement adapted to the spacing of the hexamers of the cognate HREs. Only in this configuration is a high affinity, cooperative binding possible. For a correct binding it is necessary that both recognition helices optimally contact both repeat sequences. If, as a result of an incorrect spacing of the hexamers, only one of the two recognition sequences bind, then a high affinity complex cannot be formed.

4.3.4 Ligand-binding Domains

The region E with the ligand-binding domain harbors three important functions:
- homo- and heterodimerization
- binding of ligand, both agonists and antagonists
- transactivation and transrepression: binding of coactivators and corepressors.

Dimerization

A contribution to the dimerization of the receptors – in addition to that from the DNA-binding domain – is provided by a dimerization element in the ligand-binding domain. The structure of the ligand-binding domain of RXR without bound hormone shows a homodimer with a symmetric dimerization surface, formed essentially from two antiparallel α-helices (Bourguet et al., 1995). The folding of the dimerization surface is supported by the periodic arrangement of hydrophobic amino acids with a 7-residue spacing, similar to the configuration of leucine residues found in leucine zippers (see Section 1.2). These hydrophobic amino acids, however, do not participate directly in the dimerization.

Ligand Binding

The crystal structures of the ligand-binding domains of several nuclear receptors have been resolved demonstrating a similar overall structure (review: Weatherman et al., 1999). Fig. 4.8a shows the receptor for *all-trans* retinoic acid, RAR, with is bound ligand. The ligand-binding domain is formed from 12 α-helices numbered from H1 to H12. In the bottom half of the structure, a ligand-binding pocket is found, which accommodates the ligand. The pocket is mainly hydrophobic and of variable size for different receptors. As illustrated in Fig. 4.8b for the binding of RAR bound to all-trans-retinoic acid, contacts between the ligand and the pocket can be very extensive and include many hydrophobic contacts as well as hydrogen bonds to the polar parts of the ligand.

The size of the ligand-binding pocket can vary considerably among different receptors. Receptors with specific ligands, e.g., the T_3R, bind the ligand in a small pocket that is tailored to bind only the specific ligand, the T_3 hormone, and only small deviations from the T_3 structure are tolerated. In contrast, the ligand-binding pocket of the peroxisome proliferator activator receptor, PPAR, is of a much larger size. This receptor binds a large variety of endogenous ligands like fatty acids with rather low affinity. Apparently, the ligand-binding cavity of PPAR has been adopted to bind hydrophobic ligands of different size.

Ligand binding induces a considerable conformational change in the ligand-binding domain. The liganded structures are more compact than the unliganded ones, and a reorganization of distinct α-helices is visible. A major conformational change is seen for the H12 helix that reorients by a mechanism like a "mouse trap" (review: Steimetz et al., 2001).The H12 helix is amphipathic, possessing a hydrophobic and a hydrophilic face. In the unliganded RXR, helix 12 projects away from the body of the ligand-binding domain. In the liganded structure, the helix reorients (Fig. 4.9). The hydrophobic residues face inward and form part of the ligand-binding pocket, whereas the conserved polar residues face outward and form a binding surface for the binding of coactivators and corepressors. Both structural data and mutational analyses demonstrate that the H12 helix is directly involved in transcriptional activation. It forms the core of the AF2 domain and thus plays a key role in transactivation and transrepression.

The nuclear receptors are involved in the regulation of numerous physiological processes and are therefore important medical targets. An enormous number of synthetic

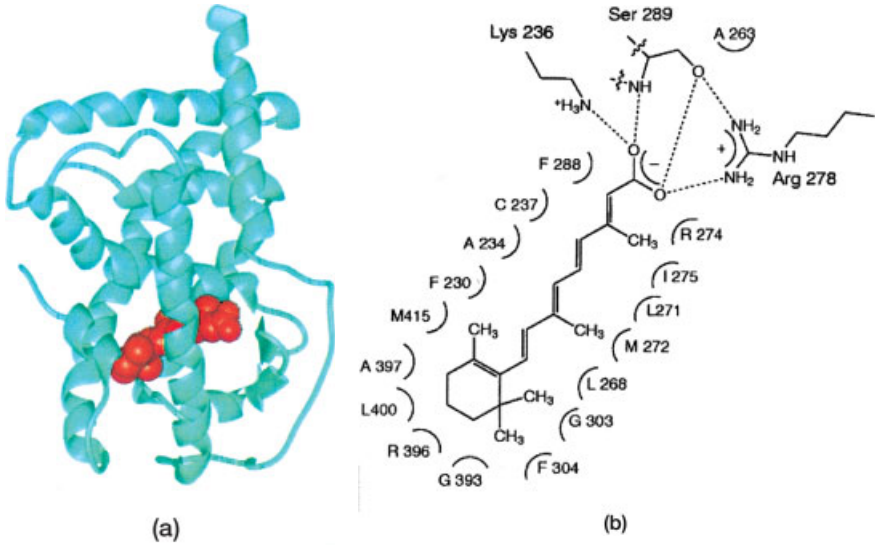


Fig. 4.8 a) Ribbon diagram of RAR bound to all-*trans* retinoic acid. b) Schematic diagram of retinoic acid binding site. H-bonds and ionic interactions fix the carboxylate of the ligand whose nonpolar parts are deeply buried in a hydrophobic pocket.

compounds are available that act either as agonist or as antagonist of the natural ligands. As illustrated in Fig. 4.10, the structure of a ligand-binding domain with an antagonist bound can provide a rational basis to explain the antagonistic function of a ligand. In the presence of the antagonist, the conformational change within the ligand-binding domain is quite distinct from that induced by the natural ligand. Helix H12, which is critical for transactivation, assumes a conformation distinct from that observed for the natural ligand. The structure suggests that the binding surface on H12 that is required for coactivator binding is not available in the antagonist-bound state, thereby preventing transcriptional activation.

4.3.5 Transactivating Elements of the Nuclear Receptors

Most nuclear receptors contain two structural elements, AF-1 and AF-2, which mediate the transcriptional activation. The AF-1 domain is located in the A/B region and mediates a *ligand-independent transactivation*. It harbors phosphorylation sites and interaction sites for coactivators. Overall, the role of the AF-1 domain in transcriptional activation is not well understood. The AF-2 domain of the E region is well conserved across the members of the nuclear receptor superfamily and functions in a ligand-dependent way. It encompasses helix H12 which is part of the ligand-binding domain as well as other structural elements of the E region. Upon ligand binding, the hydrophilic surface of the H12 helix is oriented outward, providing an interface for the binding of complementary regions of coactivators or corepressors.

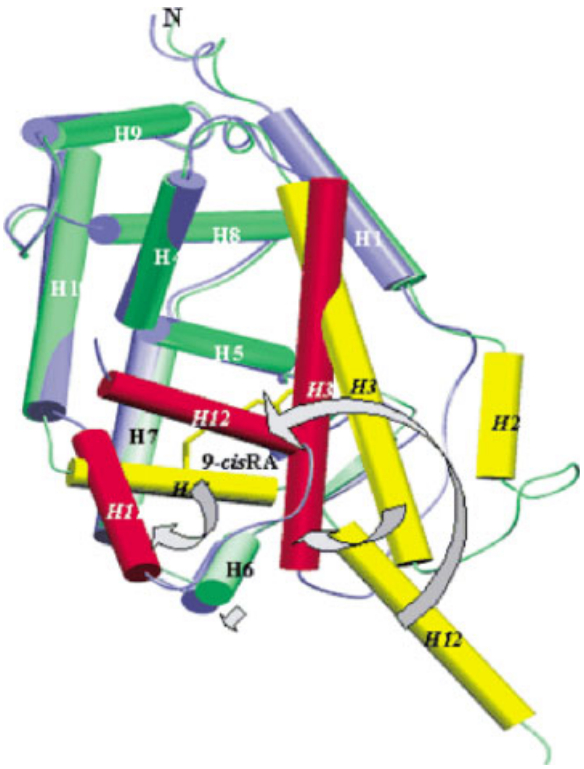


Fig. 4.9 Structural changes in the ligand binding domain of RXR on binding of 9-*cis* retinoic acid. The models of domain E of apo-RXRα and the binary complex of RXRα and 9-*cis* retinoic acid were superimposed. Domain E of apo-RXRα is depicted in green and yellow, and domain E of the binary complex in blue and red. The grey arrows indicate the structural rearrangements of helices 11, 12, and the N-terminus of helix 3. After Steinmetz et al., 2001.

4.4 Mechanisms of Transcriptional Regulation by Nuclear Receptors

Most of the functions of nuclear receptors can be described in terms of activation and repression of transcription. Although attention has been focussed primarily on the transcription activation mediated by "positive" HREs, it is increasingly recognized that nuclear receptors can also repress transcription in a ligand-dependent manner. So-called "negative HREs" have been identified that bind the receptor and mediate negative regulation by the ligand. These elements have been identified for glucocorticoid receptors and for the T3 receptor. In addition to ligand-dependent activation and repression, a subset of nuclear receptors repress basal transcription in the absence of ligand when bound to a positive HRE, thereby silencing the target gene (Fig. 4.11).

Multiple protein-protein interactions are involved in the transcription regulation by nuclear receptors. Nuclear receptors have been shown to directly interact with

- components of the basal transcription apparatus, e.g., TFIIB and TAF proteins
- transcriptional coactivators
- transcriptional corepressors
- other specific transcription factors.

Most detailed information is available for the binding of coactivators and corepressors.

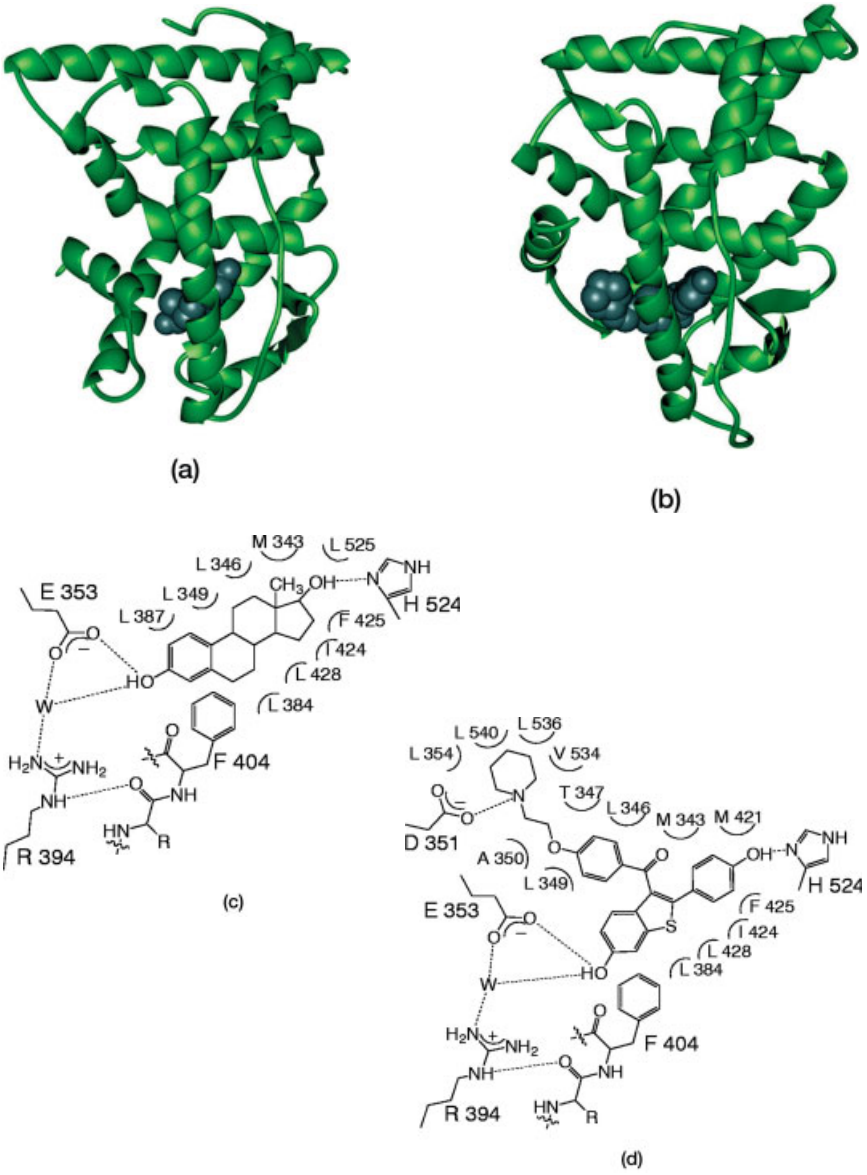


Fig. 4.10 Agonist vs antagonist binding to estrogen receptor, ER
a) Ribbon diagram of ER bound to the agonist estradiol (black). b) Ribbon diagram of ER bound to the antagonist raloxifene (black). c) schematic drawing of estradiol-binding site. d) schematic diagram of raloxifene-binding site.

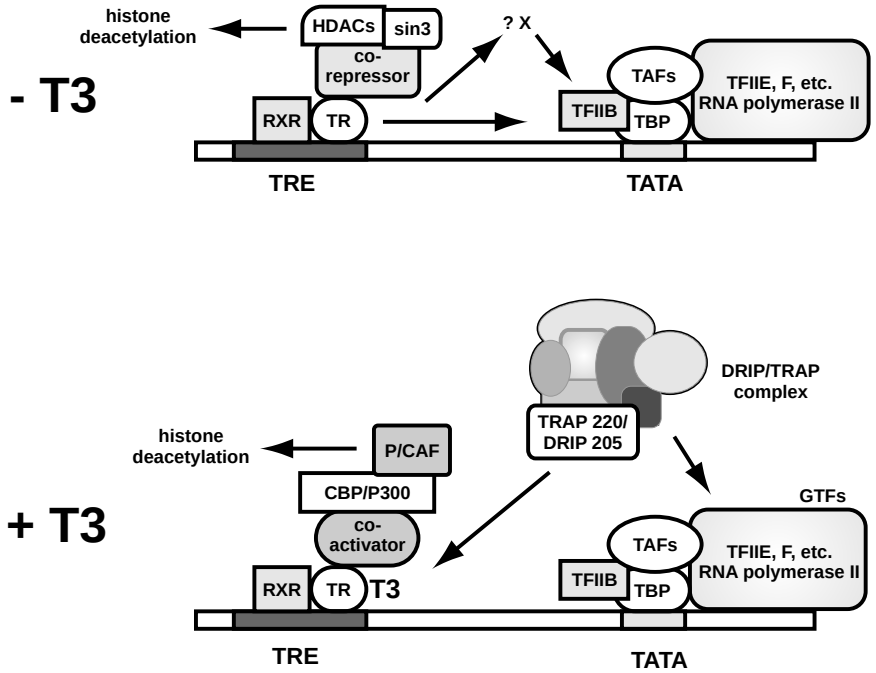


Fig. 4.11 Model of repression and activation of T_3R. In the absence of the T3 hormone, a heterodimeric RXR-T_3R receptor is bound at the T3-responsive element,TRE, establishing a basal repressed state. The repressed state is maintained by recruitment of corepressor complexes containing histone deacetylase activity. X refers to potential unidentified cofactors (possibly chromatin remodeling complexes or SRBs) which help to keep the promotor-bound basal transcription apparatus in the inactive state. In the presence of T_3 hormone, the corepressors are removed and coactivators (e.g. the SRC/p160 complex) bind to the receptor heterodimer. The histone acetylase activity of the associated proteins helps to induce a transcription competent state of the nucleosomes. Furthermore, another coactivator complex, DRIP/TRAP, is recruited by to the promotor which is thought to stabilize binding of the RNA polymerase holoenzyme. After Yen (2001).

Coactivators of Nuclear Receptors

A series of proteins or protein complexes with coactivator function for nuclear receptors has been identified that specifically interact with the activated, liganded receptor. The most abundant ones are now included in the p160 family of coactivators with the steroid receptor coactivator 1, SRC-1, as a well-characterized member. Another group of coactivators is present as multiprotein complexes like the TRAP complex (TRAP, thyroid hormone receptor activating protein). A common receptor interaction motif LXXLL is found on these coactivators which mediates at least part of the interaction with the AF2 domain of the receptor.

The p160 coactivators themselves interact with other proteins or protein complexes with mediator function and histone acetylase activity. Particularly, an interaction with the CBP/p300 protein and the PCAF histone acetylase (see Section 1.4.6.) has been shown. It is thought that the primary function of the p160 class of coactivators is

to concentrate HAT activity at target gene promotors and that the role of the TRAP proteins is to establish a link between the DNA-bound receptor and the RNA polymerase II holoenzyme.

Other types of coactivators include the CBP/p300 proteins themselves, which can also directly interact with DNA-bound receptors in a ligand-dependent manner. Members of the SWi/SNF complex and the E3 ubiquitin protein ligase E6-AP have also been recognized as having the function of coactivators of nuclear receptor function (review: McKenna et al., 1999).

For many specific transcriptional activators, the recruitment of histone acetylase activity and a subsequent reorganization of chromatin has been identified as an essential step in transcription activation of chromatin-covered DNA (see Section 1.4.6). The identification of nuclear receptor coactivators with direct or associated histone acetylase activity underscores that nuclear receptors use the recruitment of histone acetylase as an essential step in transcription activation as well. The association of various various histone acetylase activites with the coactivators of the nuclear receptors indicates that distinct protein complexes with HAT activity can be recruited to DNA-bound nuclear receptors. It is still an open issue which specific functions are performed by the different HAT activities during transcription activation.

Corepressors of Nuclear Receptors

Unliganded T_3receptor (T_3R) and retinoic acid receptor (RAR) can repress transcription in the absence of ligand. This transcriptional repression is linked to the binding of proteins with corepressor activity (review: Zhang and Lazar, 2000). Examples of such corepressors are the nuclear corepressor NcoR and the "silencing mediator for retinoic and thyroid hormone receptors", SMRT. Both proteins interact with the ligand-binding domain and appear to be released from the receptor upon ligand binding. Possibly alternative conformations of the AF-2 region serve to trigger this release.

A central tool for gene repression is the recruitment of histone deacetylase activity to the promotor region (see Section 1.4.6). Both the NCoR and SMRT proteins have been shown to interact with protein complexes like the mSin3 complex and the NuRD complex that contain histone deacetylase activity. This observation strongly suggests that the nuclear receptor corepressors act by recruiting histone deactylase activity to the promotor region, thereby maintaining a repressed state of chromatin.

A schematic representation of the current model of coactivator and corepressor function in nuclear receptor signaling is shown in Fig. 4.11.

4.5
Regulation and Variability of Signaling by Nuclear Receptors

Signaling by nuclear receptors is regulated at various levels (Fig. 4.12). The following are important regulatory attack points:

- **Regulation at the level of ligand concentration**
  A main determinant of nuclear receptor signaling is the concentration of the ligand available for binding. The ligand concentration can be regulated in many ways (for details see textbooks on hormone action):
  – synthesis and degradation
  – modification
  – secretion, transport and storage
  – feedback regulation via the circulating hormone concentration.
  The latter point is of importance in the hypothalamus-pituitary system of the brain, where feedback inhibition at various levels is used to prevent overproduction of, e.g., steroid hormones or the T3 hormone.
- **Crosstalk: regulation by phosphorylation**
  The phosphorylation of nuclear receptors on Ser/Thr residues is a key mechanism for the coupling of nuclear receptor signaling to other signaling pathways of the cell. Many nuclear receptors are isolated as phosphoproteins, and their phosphorylation provides a means for ligand-independent activation and regulation. The steroid

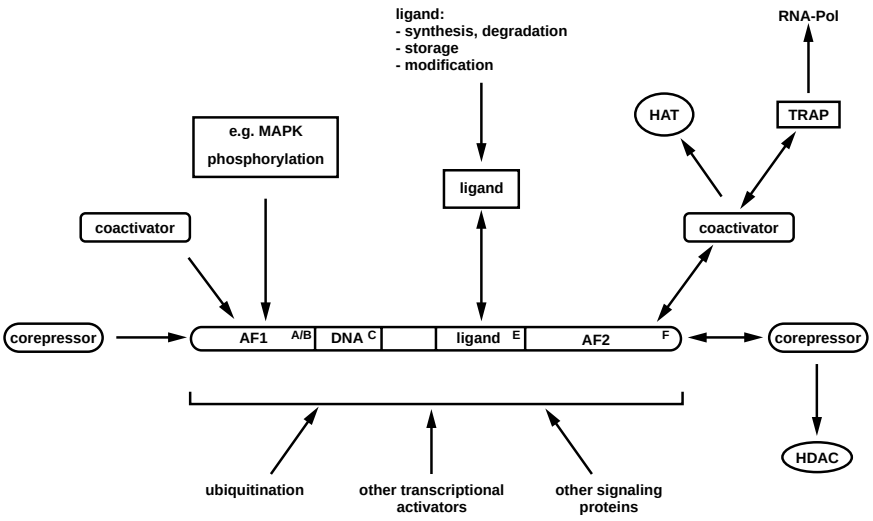


Fig. 4.12 Functions of nuclear receptor domains. The domains A/B, C, E and F of the nuclear receptors are involved in multiple protein-protein interactions and are subject to regulatory modifications as indicated. Most important are the corepressor and coactivator complexes that direct histone deacetlyase (HDAC) and histone acteylase (HAT) activities, respectively, to the nuclear receptor regulated promotor region. TRAP: thyroid hormone receptor activating protein.

hormone receptors are phophoproteins which are usually phosphorylated on several positions. The phosphorylation sites are mainly found in the N-terminal region of the receptors. The consequences of phosphorylation for the receptor proteins are varied. It is conceivable, and in some cases experimentally proven, that it has influence on hormone binding, nuclear transport, DNA binding and transactivation. The estrogen receptor is phosphorylated on Ser 111 of its A/B region by kinases of the MAPK cascade (see Chapter 10). This phosphorylation results in ligand-independent activation of the receptor and allows for the transcription of ER-target genes in the absence of estrogen. The MAPK cascade transmits signals from growth factor-stimulated pathways and from Ras proteins (see Chapter 9), providing a crosstalk between nuclear receptor signaling and growth factor signaling. Other signaling pathways like the protein kinase A pathway (see Section 7.3), the TGFβ-pathway (see Chapter 12) and the Akt kinase pathway (see Section 6.6.3) have been shown to influence nuclear receptors by phosphorylation as well. Androgen receptor-mediated apoptosis is, e.g., suppressed upon phosphorylation of the androgen receptor by Akt kinase. The examples illustrate that growth-inhibiting and growth-promoting signals can be transmitted through nuclear receptors, indicating extensive crosstalk between nuclear receptor signaling and other central signaling pathways. The signals transmitted by nuclear receptors are thus integrated into the intracellular signaling network and are linked to extracellular signals other than the hormone ligands and to protein kinase cascades. The detailed structural and functional consequences of nuclear receptor phosphorylation remain, however, to be elucidated.

In addition to the receptors themselves, the nuclear receptor coactivators have also been found to be targets of regulatory phosphorylation events.

- **Interaction with other transcriptional activators**

  Nuclear receptors can also modulate gene expression by interference with the activity of other transcriptional activators. The ER, T_3R, RAR and GR proteins have, e.g., been shown to act as transrepressors of the transcription factor AP-1, which is a heterodimer composed of c-Jun and c-Fos proteins. Reciprocally, AP-1 can inhibit transactivation by these receptors. A mutually antagonistic effect of glucocorticoid receptors on the function of the transcriptional activator NFκB has also been reported. Furthermore, a mutual interference between different nuclear receptors has been demonstrated. In the reproductive system, progesterone, acting through the progesterone receptor, is a physiological negative regulator of the estrogen receptor *a*. It is postulated, that the activated progesterone receptor prevents the estrogen receptor from activating transcription (review: McDonnell and Norris, 2002). The mechanistic basis of this antagonism is not understood. Competition for common coactivators or mediators may be involved.

- **Regulation by ubiquitination**

  As for other central regulators, the level of nuclear receptors is modulated by ubiquitin-mediated degradation (review: Dennis and Nawaz, 2001). A ligand-dependent ubiqutination and subsequent proteasomal degradation has been described for ER*a*, PR, VDR, T_3R and RAR*a*, allowing for a down-regulation of signaling by these receptors. In addition, corepressors and coactivators of nuclear receptors

are also subjected to ubiquitination indicating a widespread involvement of the ubiquitin-proteasome pathway in the regulatory function of nuclear receptors.

Overall, the nuclear receptors are engaged in a large number of protein-protein interactions. For the estrogen receptor, a multitude of contacts and links to other regulatory proteins have been established. Most of these proteins interact with different target proteins linking the receptor to other signal transduction pathways. In this way, the estrogen receptor is embedded in a signaling network that allows for tissue- and cell-specific regulation of estrogen action (review: Mc Donnell and Norris, 2002).

4.6 The Signaling Pathway of the Steroid Hormone Receptors

Based on the receptor activation mechanism, the nuclear receptors may be divided into two basic groups. In the first group (those including most of the steroid hormone receptors), the receptors can be localized in the nucleus or in the cytoplasm. The receptors of the other group (discussed in Section 4.7) are always localized in the nucleus. Representative ligands of these receptors are the derivatives of retinoic acid, the T_3 hormone and $VitD_3$.

Signal transduction by steroid hormones is distinguished by the fact that the receptors can be found either in the cytoplasm or in the nucleus. The steroid hormone receptor receives the hormonal signal in the cytosol, becomes activated by hormone binding, at which point it enters the nucleus to regulate the transcription initiation of cognate genes. Figure 4.13 shows the most important steps in the signal transduction by steroid hormones.

Activation of the Cytoplasmic Aporeceptor Complexes

The steroid hormones are distributed throughout the entire organism by means of the circulatory system. The transport often occurs in the form of a complex with a specific binding protein. An example of such a binding protein is transcortin, which is responsible for the transport of the corticosteroids. The steroid hormones enter the cell by diffusion and activate the cytosolic receptors.

In the absence of steroid hormones, the receptors remain in an inactive complex, designated the *aporeceptor complex*. In the aporeceptor complex the receptor is bound to proteins belonging to the chaperone class. Chaperones are proteins whose levels are increased as a result of a stress situation, such as a rise in ambient temperature. The chaperones assume a central function in the folding process of proteins in the cell. Chaperones aid proteins in avoiding incorrectly folded states. They participate in the folding of proteins during and after ribosomal protein biosynthesis, during membrane transport of proteins, and in the correct assembly of protein complexes.

With the help of co-immunoprecipitation it could be shown that the receptors of steroid hormones interact with a multi-protein chaperone machinery, including the proteins *Hsp90, Hsp70* and *Hsp56* (Fig. 4.13). The term "Hsp" (Heat shock protein) is derived from the observation that these proteins were produced at higher levels

following heat treatment. Furthermore, one finds a 23 kDa acidic protein in the aporeceptor complex whose function is not yet clear.

It is assumed that the receptor, when complexed with Hsp90, is fixed in an optimal conformation for hormone binding. The activation of the receptor by the hormone

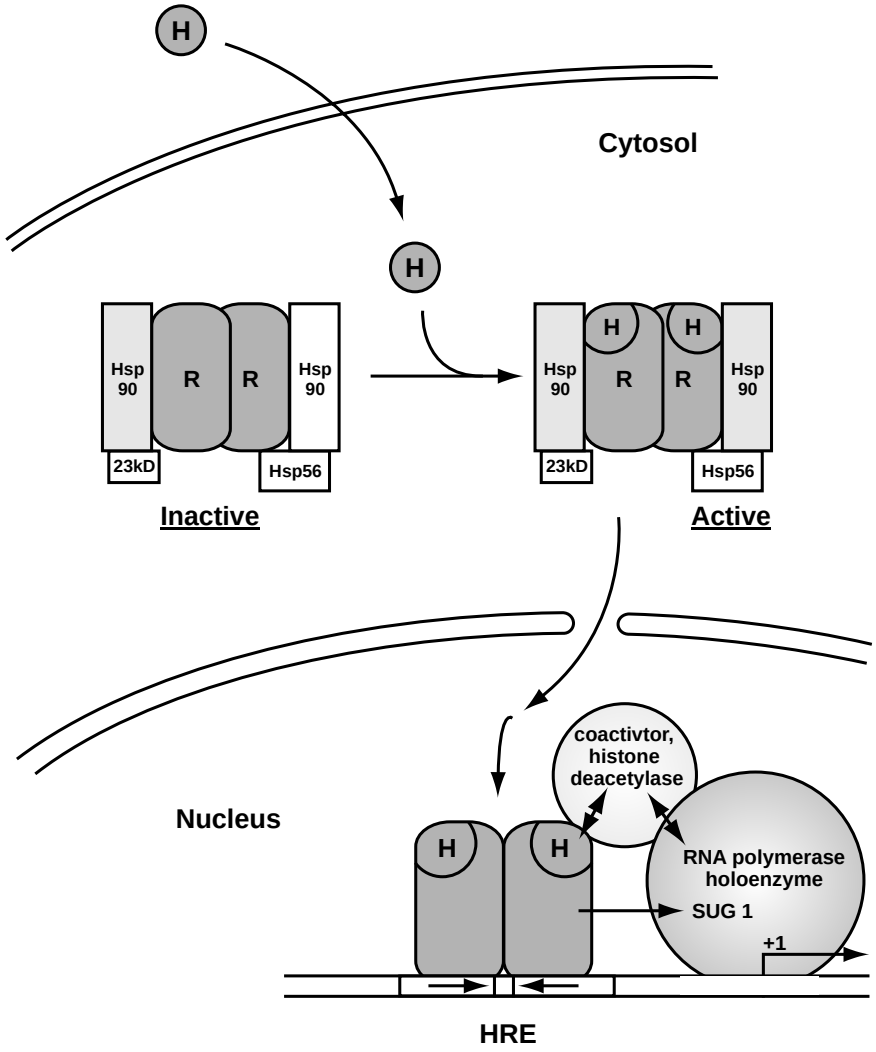


Fig. 4.13 Principle of signal transduction by steroid hormone receptors. The steroid hormone receptors in the cytosol are found in the form of an inactive complex with the heat shock proteins Hsp90 and Hsp56 and with protein p23. The binding of the hormone activates the receptor so that it can be transported into the nucleus where it can bind to its cognate HRE. It remains unclear in which form the receptor is transported into the nucleus, and to which extent the associated proteins are involved in the transport. One mechanism of activation of transcription initiation involves mediation by the proteins RIP-140 and Sug1.

requires the presence of Hsp90, which binds the receptor as a dimer. Hsp90 binds preferentially on partially unfolded proteins, and it is believed that Hsp90 holds the receptor in a partially unfolded conformation that is competent for ligand binding. It is assumed that chaperones of the Hsp70 class facilitate the correct association of Hsp90 with the receptor. Other co-chaperones like Hsp 56 and p23 participate in the activation as well. The binding of the hormone to the aporeceptor complex leads to activation of the receptor and initiates the translocation of the receptor into the nucleus. The activated receptor possesses an accessible nuclear localization sequence and is furthermore capable of DNA binding and transactivation. The ability to transactivate implies that the transactivating domain is properly positioned, as a result of the hormone binding, to allow stimulatory interactions with the transcription apparatus (see Section 4.4).

Details of the activation and transport into the nucleus remain unknown. Additional proteins like RAP46 have been identified that function in cooperation with the chaperones and participate in nuclear transport and also activation of the receptor. After dissociation of the heat shock proteins, the hormone-receptor complex is capable of specific binding on the HRE and of transactivation.

The chaperones are used as tools in this system for regulation of activity of the steroid hormone receptors. The function of the chaperones is obviously to fix the receptor in a conformation which allows high affinity binding to the hormone and the subsequent steps of specific DNA binding and transactivation. For the steroid hormone receptors this means that they must exist in functionally different conformations. It may be a function of the chaperones to stabilize the particular conformation optimal for hormone binding. Inactive, chaperone-bound nuclear receptor complexes can be also found in the nucleus, as is the case for the estrogen receptor. The estrogen receptor is sequestered in the nucleus within a large inhibitory heat shock protein complex. Binding of estrogen to the receptor enables the displacement of the heat shock proteins and facilitates DNA binding, formation of receptor dimers and interactions with coactivators and the transcription apparatus.

4.7
Signaling by Retinoids, Vitamin D3, and the T3-Hormone

Ligands of the RXR-heterodimer group and the "orphan" receptors are chemically more diverse than the ligands of the steroid family. Representative ligands of this group are the retinoids all-trans retinoic acid, 9-cis retinoic acid, the T_3 hormone and vitamin D_3 (Fig. 4.1).

In contrast to signal transduction by the steroid hormone receptors, there are multiple pathways by which the ligands of this group are made available for receptor activation (Fig. 4.14):

- The hormone ligands can be secreted in the classical endocrinological pathway and transported to the target cell where they bind the receptor.

- The active hormones can be formed intracellularly from inactive precursors. The inactive precursor is transported through the bloodstream to the target cell where it is enzymatically converted to the active hormone. An example of this pathway is that of 9-cis retinoic acid, which is synthesized from the alcohol of vitamin A (vitamin A1; retinol).
- The hormone is synthesized in the same cell in which receptor activation takes place. The ligand is synthesized intracellularly and is not secreted. An example of such an intracellular hormone is prostaglandin J_2.

A further, more dramatic difference to the steroid hormone receptors is the localization of the receptors. The receptors for the retinoids (RAR and RXR, see Table 4.1), the T_3 hormone (T_3R) and vitamin D_3 (VDR) are mainly localized in the nucleus, and their activity is not controlled by the heat shock proteins. The receptors also bind the corresponding HRE in the absence of hormone, in which case they can then act as *repressors* of gene activity. In the presence of the hormone an activation of gene expression is usually observed. However we also know of rare examples where binding of the ligand has an inhibitory effect on gene activation.

As already illustrated in Fig. 4.5, the receptors of all-trans retinoic acid, 9-cis retinoic acid, vitamin D_3 and the T_3-hormone usually bind their HREs preferentially in the form heterodimers.

RXR plays a special role in the formation of heterodimers: the receptor for 9-cis retinoic acid is usually one of the binding partners in the heterodimer.

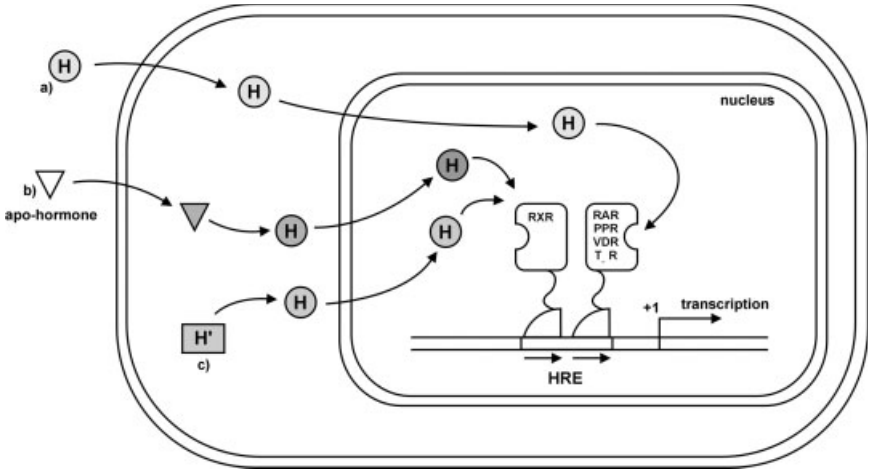


Fig. 4.14 Principle of signal transduction by RXR heterodimers. The activated hormone can be made available to the RXR heterodimer in three different ways. a) The hormone (e.g. T_3 hormone) is synthesized in endocrinal tissue and reaches the DNA-bound RXR heterodimer in the nucleus via passive transport. b) The active hormone is formed in the cytosol from an inactive apo-hormone (as for, e.g. 9-*cis*-retinoic acid). c) The hormone is synthesized intracellularly. In all three cases, the binding of the hormone-RXR-heterodimeric complex is the signal that induces transcription activation of the downstream genes. After Mangelsdorf and Evans, 1995.

4.7.1 Structure of the HREs of RXR Heterodimers

The HREs of RXR-heterodimers are usually composed of two identical or nearly identical copies of the hexamer sequence AGGTCA in direct repeat. The apparently simple structure of the HREs leads to the question of how the receptors of this class can distinguish between the various HREs. Studies with artificial HRE constructs, as well as of naturally occurring HREs, indicate a complicated cooperative effect between HRE structure on the one hand and homo- or heterodimer formation of the receptors on the other hand.

The following points were identified as important for the recognition and discrimination of a particular HRE:

- In the case of identical hexamer sequences of an HRE, the spacing between the hexamers is a specificity-determining element *(n-rule)*. The spacing can be between 1 and 6 bp. Grounds for the discrimination based on spacing is the structure of the receptor dimer. A given receptor demands a particular spacing of the hexamers in the HRE because of steric requirements. If the distance requirement is not met, the high affinity and cooperative formation of a dimer on the HRE is not possible.
- The receptor for 9-cis retinoic acid (RXR) usually occupies the 5' position in the heterodimer. The RXR serves as a quasi vehicle to bring other receptor monomers to the 3' half-site of the HRE.
- Apart from the spacing, the sequence of the hexamers is also important for the discrimination. In naturally occurring HREs, one finds small deviations from the consensus sequence AGGTCA. It has been shown that the exact nature of the HRE sequence as well as neighboring sequences contribute to the identity and specificity of an HRE.
- for heterodimers out of VDR and T_3R it may be shown that the hormone concentration necessary for activation depends on which of the two monomers occupies the 5' and which the 3' position in the HRE.

4.7.2 Complexity of the Interaction between HRE, Receptor and Hormone

There is an altogether complex interaction between HRE, receptor and hormone in the group of receptors for retinoids, vitamin D_3 and the T_3-hormone. The complexity is determined by the following factors:

- **Formation of the Homo- or Heterodimer**
  To what extent hetero- or homodimers are formed depends upon the concentration of the various receptors in the particular cell type. The receptor concentration can furthermore be regulated via the general mechanisms as described in Chapters 1 and 2, whereby regulation at the level of expression of the receptors is particularly important.

- **Multiplicity of the HREs**
  The identity of an HRE is determined by sequence, spacing and polarity of the hexamer sequences. There are many possible combinations of these specificity determinants, such that the hormonal activation of a gene decidedly depends on the nature of the cognate HRE. There is an additional, poorly understood influence from the neighboring sequences, and, furthermore, HREs are often arranged in tandem. Another factor which can influence HRE activity is the occurrence of several different regulatory elements on one promoter. Further DNA elements are often found in the neighborhood of the HRE, which can serve as further transcriptional activators, such as, for example, AP1. In this manner, other transcriptional activators can act cooperatively with or antagonistically against the receptor.
- **Multiplicity of the Receptors**
  Among the receptors for the retinoids (RAR and RXR), there are at least three subtypes characterized, known as RARα, β, and γ (RXR α,β and γ). The various subtypes differ from each other mainly in their amino acid sequence and are encoded in each case by their own gene. Furthermore, one finds isoforms of some of the receptors created by alternative splicing of the primary transcript. Altogether, the repertoire of the various receptor variants is enormously increased through the existence of the subtypes and by alternatively spliced receptors. Because the individual variants differ in their ability to form hetero- and homodimers as well as in their ability to activate and bind ligands, this results in a great variety of functionally different receptors.
- **Binding and Activation via Hormones**
  With regard to ligand binding, heterodimer formation brings a further element of complexity into the picture. The ligand concentration necessary for activation varies strongly depending upon the nature of the heterodimer. For example, there are RXR-heterodimers, such as RXR-T_3R, in which the RXR binding site for 9-cis retinoic acid is not accessible. In this case, the RXR is a "silent" partner. In other combinations (e.g., RXR-PPAR), the ligands of both receptors are required for full activation of the heterodimers.

The hormone concentration necessary for the binding and activation of a receptor dimer can be very different for the various dimers. The available hormone concentration in a particular cell type thus plays an important role in receptor activation. The effective intracellular concentration of a hormone is furthermore subject to a diverse regulation mechanism (see Section 4.5).

The picture sketched above for the function of the receptors of the retinoids and the T_3- and vitamin D_3-hormones is in many points still incomplete. The cooperation of the activated receptors with other transcription factors, their interactions with the transcription apparatus and the influence of chromatin structure are still poorly understood.

Reference

1 Aranda, A. and Pascual, A. (2001) Nuclear hormone receptors and gene expression. *Physiol Rev.*, **81**, 1269–1304.

2 Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. and Moras, D. (1995) Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-alpha. *Nature*, **375**, 377–382.

3 Dennis, A.P., Haq, R.U. and Nawaz, Z. (2001) Importance of the regulation of nuclear receptor degradation. *Front Biosci.*, **6**, D954–D959.

4 Kraulis, P.J. (1991) 'MOLSKRIPT: A program to produce both detailed and schematic plots of protein structures' J. Appl. Crystallogr. **24**, 946–950

5 Lin, H.M., Zhao, L. and Cheng, S.Y. (2002) Cyclin D1 Is a Ligand-independent Co-repressor for Thyroid Hormone Receptors. *J.Biol.Chem.*, **277**, 28733–28741.

6 Mangelsdorf, D.J. and Evans, R.M. (1995) The RXR heterodimers and orphan receptors. *Cell*, **83**, 841–850.

7 McDonnell, D.P. and Norris, J.D. (2002) Connections and regulation of the human estrogen receptor. *Science*, **296**, 1642–1644.

8 McKenna, N.J., Xu, J., Nawaz, Z., Tsai, S.Y., Tsai, M.J. and O'Malley, B.W. (1999) Nuclear receptor coactivators: multiple enzymes, multiple complexes, multiple functions. *J. Steroid Biochem. Mol. Biol.*, **69**, 3–12.

9 Perlmann, T and Evans, R.M. (1997) 'Nuclear receptors in Sicily: all in the famiglia' *Cell* **90**, 391–397

10 Robinson-Rechavi, M., Garcia, H.E. and Laudet, V. (2003) The nuclear receptor superfamily. *J. Cell Sci.*, **116**, 585–586.

11 Steinmetz, A.C., Renaud, J.P. and Moras, D. (2001) Binding of ligands and activation of transcription by nuclear receptors. *Annu. Rev. Biophys. Biomol. Struct.*, **30**, 329–359.

12 Weatherman, R.V., Fletterick, R.J. and Scanlan, T.S. (1999) Nuclear-receptor ligands and ligand-binding domains. *Annu. Rev. Biochem.*, **68**, 559–581.

13 Zhang, J. and Lazar, M.A. (2000) The mechanism of action of thyroid hormones. *Annu. Rev. Physiol*, **62**, 439–466.

5
G Protein-Coupled Signal Transmission Pathways

5.1
Transmembrane Receptors: General Structure and Classification

During intercellular communication, extracellular signals are registered by the cell and converted into intracellular reactions. Signal transmission into the cell interior takes place by reaction chains, in which several individual reactions generally run in sequence and involve many signal proteins. The nature of the extracellular signal can be very diverse and may include extracellular signal molecules, such as low-molecular-weight messenger substances or proteins, or sensory signals such as light signals. The cell uses two principal ways to transduce signals into the interior of the cell. One way is exemplified by nuclear receptor signaling, where the signaling molecule crosses the cell membrane and activates the receptor in the interior of the cell. In the other major way the signal is registered at the cell membrane and transduced into the cell by transmembrane proteins. Three different types of transmembrane proteins participate in this mode of signaling:

Signaling via Transmembrane Receptors

Transmembrane receptors are proteins that span the phospholipid bilayer of the cell membrane. The signaling molecule binds on the extracellular side to the receptor, which is thereby activated. Reception of the signal is synonymous with activation of the receptor for transmission of the signal across the cell membrane. Transmission of the signal implies specific communication with the effector protein, the next component of the signal transmission pathway on the inner side of the cell membrane. In this process enzymatic activities can be triggered and/or the activated receptor engages in specific interactions with downstream signal proteins. An *intracellular signal chain* is set in motion, which finally triggers a defined biochemical response of the target cell (Fig. 5.1a). Sensory signals (light, pressure, odor, taste) can be received as well by transmembrane receptors and can be transmitted into intracellular signals.

An example of a transmembrane receptor that registers sensory signals is rhodopsin. Rhodopsin is a sensory receptor that plays a role in vision by receiving light signals and converting them into intracellular signals.

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

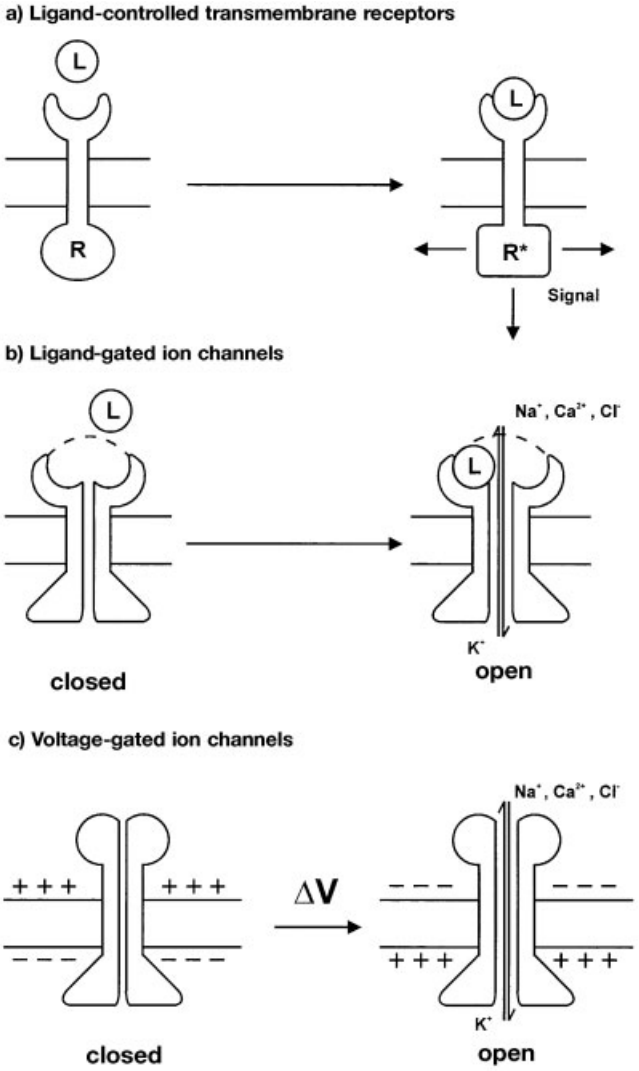


Fig. 5.1 Mechanism of signal transduction at membranes. a) Signal transmission via ligand-controlled transmembrane receptors. The ligand L binds to the extracellular domain of a transmembrane receptor, whereby the receptor is activated for signal transmission to the cytosolic side. The cytosolic domain of the activated receptor R* transmits the signal to signal proteins next in sequence. b) Signal transduction via ligand-controlled ion channels. The ligand binds to the extracellular side of a receptor that also functions as an ion channel. Ligand binding induces the opening of the ion channel, there is an ion efflux and a change in the membrane potential. c) Signal transduction via voltage-gated ion channels. A change in the membrane potiential V is registered by an ion channel which transitions from the closed to the open state.

Signaling via Ligand-gated Ion Channels

One simply designed path of signal transmission is found in neuronal communication. Transmembrane receptors are also used for signal transmission here. These have the character of a *ligand-gated ion channel* (Fig. 5.1b). Binding of a ligand (neurotransmitter or neurohormone) to the transmembrane receptor leads to a conformational change of the receptor that enables the flow of ions through the membrane. In this case, the receptor presents itself as an ion channel with an open state controlled by ligand binding to the outer side (or also to the inner side).

Another mechanism of signaling across the cell membrane uses changes in membrane potential. A change in membrane potential induces the opening of an ion channel, and ions cross the membrane. In this case, the change of the ion's milieu is the intracellular signal. Ion channels with an open state regulated by changes in membrane potential are known as *voltage-gated ion channels* (Fig. 5.1c). The potential-driven passage of ions through ion channels is the basis for stimulation in nerves.

Intracellular Activation of Receptors

We also know of transmembrane receptors for which the reception of the signal and activation take place on the *inner side* of the membrane. The cGMP-dependent ion channels involved in signal conduction in the vision process are ligand-regulated ion channels with an open state controlled by intracellularly created cGMP. Another example is the receptors for inositol triphosphate which are localized in the membrane of Ca^{2+} storage organelles and also have the character of ligand-controlled ion channels. Inositol triphosphate is an intracellular messenger substance that binds to the cytosolic side of the corresponding receptor located in the membrane of cell organelles. Ligand binding leads to opening of the ion channel via a conformational change and thus to influx of Ca^{2+} ions from the storage organelle into the cytosol (see Section 6.5).

5.2 Structural Principles of Transmembrane Receptors

Transmembrane receptors are *integral membrane proteins*, i.e., they possess a structural portion that spans the membrane. An *extracellular domain*, a *transmembrane domain* and an *intracellular* or *cytosolic domain* can be differentiated within the structure (Fig. 5.2a).

5.2.1 The Extracellular Domain of Transmembrane Receptors

In many receptors, the extracellular domain contains the ligand-binding site. Glycosylation sites, i.e. attachment sites for carbohydrate residues, are also located nearby in the extracellular domain.

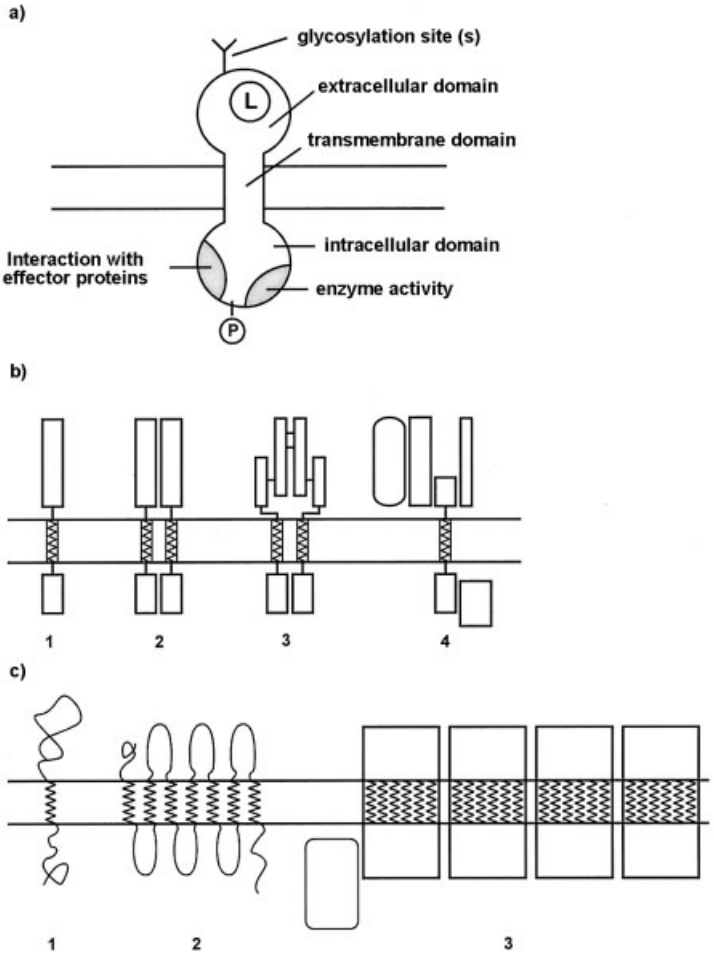


Fig. 5.2 Structural principles of transmembrane receptors. a) Representation of the most important functional domains of transmembrane receptors. b) Examples of subunit structures. Transmembrane receptors can exist in a monomeric form (1), dimeric form (2) and as higher oligomers (3,4). Further subunits may associate at the extracellular and cytosolic domains, via disulfide bridges (3) or via non-covalent interactions (4). c) Examples of structures of the transmembrane domains of receptors. The transmembrane domain may be composed of an α-helix (1) or several α-helices linked by loops at the cytosolic and extracellular side (2). The 7-helix transmembrane receptors are a frequently occurring receptor type (see 5.3). Several subunits of a transmembrane protein may associate into an oligomeric structure (3), as is the case for voltage-controlled ion channels (e.g., K^+ channel) or for receptors with intrinsic ion channel function.

The structure of the extracellular domain can be very diverse and is determined by the number of transmembrane sections as well as the subunit structure of the receptor.

The extracellular localized protein portion may be formed from a continuous protein chain and may include several hundred amino acids. If the receptor crosses the membrane with several transmembrane segments, the extracellular domain is formed from several loops of the protein chain that may be linked by disulfide bridges.

Transmembrane receptors may show *homotrophic* composition (*identical subunits*) or *heterotrophic* composition (*different subunits*, Fig. 5.2b), so that the extracellular domain may be made up of several identical or different structural elements.

We also know of receptors in which only one subunit spans the membrane, whilst other subunits are bound to this subunit on the extracellular side via protein-protein interactions or via disulfide bridges (Fig. 5.2b and examples in Chapter 11).

5.2.2 The Transmembrane Domain

The transmembrane domains have different functions, according to the type of receptor. For ligand-controlled receptors, the function of the transmembrane domain is to pass the signal on to the cytosolic domain of the receptor. For ligand-controlled ion channels, the transmembrane portion forms an ion pore that allows selective and regulated passage of ions.

The transmembrane receptors span the ca. 5 nm thick phospholipid bilayer of the cell membrane with structural portions known as *transmembrane elements*. The inner of a phospholipid layer is hydrophobic and, correspondingly, the surface of the structural elements that come into contact with the inner of the phospholipid double layer also has hydrophobic character.

The transmembrane domain may be made up of one or many transmembrane elements. Generally, the transmembrane elements include 20 – 30 mostly hydrophobic amino acids. At the interface with aqueous medium, we often find hydrophilic amino acids in contact with the polar head groups of the phospholipids. In addition, they mediate distinct fixing of the transmembrane section in the phospholipid double layer. A sequence of 20 – 30 hydrophobic amino acids is seen as characteristic for membrane-spanning elements. This property is used in analysis of protein sequences, to predict possible transmembrane elements in so-called “hydropathy plots”.

Structure of Transmembrane Elements

High-resolution structural information about the transmembrane elements of transmembrane receptors could recently be obtained on the example of rhodopsin, the light-activated G protein coupled receptor of the vision process (Fig. 5.3). These data, together with earlier data on the structures of other transmembrane proteins (e.g., bacteriorhodopsin), have confirmed that α-helices are the principal structural building blocks of the transmembrane elements of membrane receptors. The transmembrane helices are composed of 20 – 30 hydrophobic amino acids with some polar

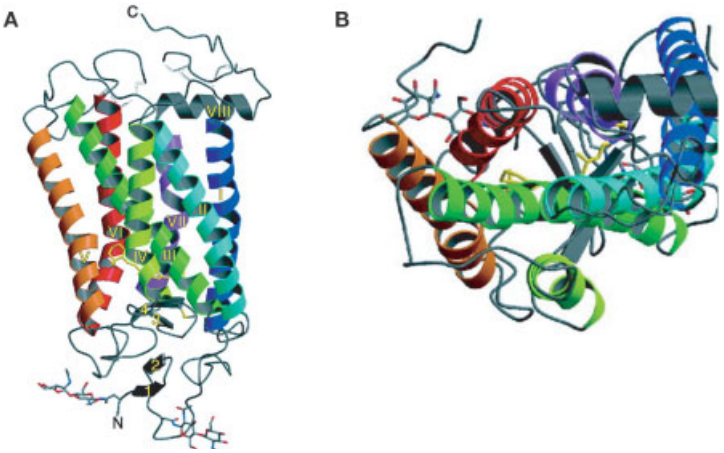


Fig. 5.3 Three-dimensional structure of rhodopsin. Two views of rhodopsin. A) The seven α-helices of the G protein-coupled receptor rhodopsin weave back and forth through the membrane lipid bilayer (yellow lines) from the extracellular environment (bottom) to the cytoplasm (top). The chromophore 11-cis retinal (yellow) is nested among the transmembrane helices .B) View into the membrane plane from the cytoplasmic side of the membrane. Roman numerals indicate numbered helices.

amino acids interspersed between or located at the helix ends. Most of the helices are arranged nearly perpendicular to the phospholopid bilayer and form bundle-like structures in which the helices are linked by loops of variable size. Ligand binding or reception of a sensory signal triggers a change in the mutual orientation of the helices that is transmitted into conformational changes of the cytoplasmic loops. These are sensed by the next components of the signaling chain or result in the activation of an enzymatic activity at the cytoplasmic side of the membrane. A signal is thereby generated in the interior of the cell and is propagated further. In addition to α-helices, proteins also use β-structures to cross the membrane. The transmembrane domain of the bacterial OmpF porin is made up of β-elements (see Fig. 5.4). The β-elements, in this case, are not mostly made up of hydrophobic amino acids and form a barrel-like structure.

Composition of the Transmembrane Domain

The transmembrane domain may consist of one or several transmembrane elements (see also Fig. 5.2). In the latter case, these are arranged in the form of bundles, as shown in Fig. 5.3 for rhodopsin. In the case of ion channels, in which several subunits are involved in the formation of the transmembrane domain, prediction of the structure of the membrane portion is very difficult. The different transmembrane elements are no longer equivalent in these cases. Part of the element is involved in formation of the inner wall of the pore; other structural elements form the surface to the hydrophobic inner of the phospholipid bilayer. It is evident that the polarity requirements for the amino acid side chains vary according to the position of the transmembrane elements (see Chapter 16).

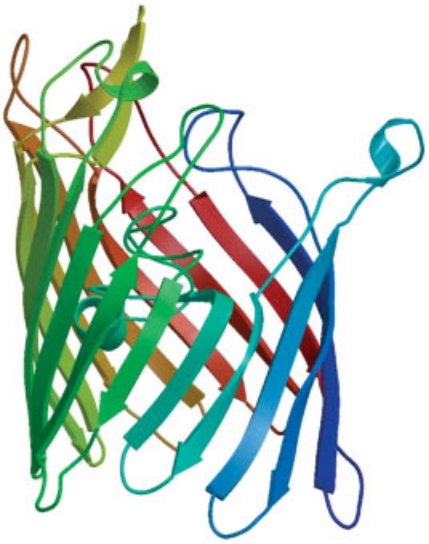


Fig. 5.4 The OmpF porin from *Eschericia coli* is an integral membrane channel-forming protein which spans the outer membrane in Gram-negative bacteria. The structure of a monomer of the OmpF porin is shown. In total, 16 β-bands are configured in the form of a cylinder and form the walls of a pore through which selective passage of ions takes place.

5.2.3 The Intracellular Domain of Membrane Receptors

Two basic mechanisms are used for conduction of the signal to the inner side of the membrane (Fig. 5.5):

- Via specific protein-protein interactions, the next protein component in the signal transmission pathway, the effector protein, is activated. The conformational change that accompanies the perception of the signal by the receptor creates a new interaction surface for proteins that are located downstream of the receptor. In the absence of the signal, this interaction surface is not available. Signal transmission therefore strictly depends on signal perception by the receptor, and activation of the effector molecule must be preceded by activation of the receptor by a signal.
- Arrival of the signal triggers enzyme activity in the cytosolic domain of the receptor, which, in turn, pulls other reactions along with it. The enzyme activity of the cytosolic domain is often tyrosine kinase activity; however, there are other examples where tyrosine phosphatase or Ser/Thr-specific protein kinase activity is activated. In all these examples, the cytoplasmic domain carries an enzyme activity regulated by ligand binding. The enzyme activity may be an integral part of the receptor, or it may also be a separate enzyme associated with the receptor on the inner side of the membrane (cf. Chapters 8 and 12).

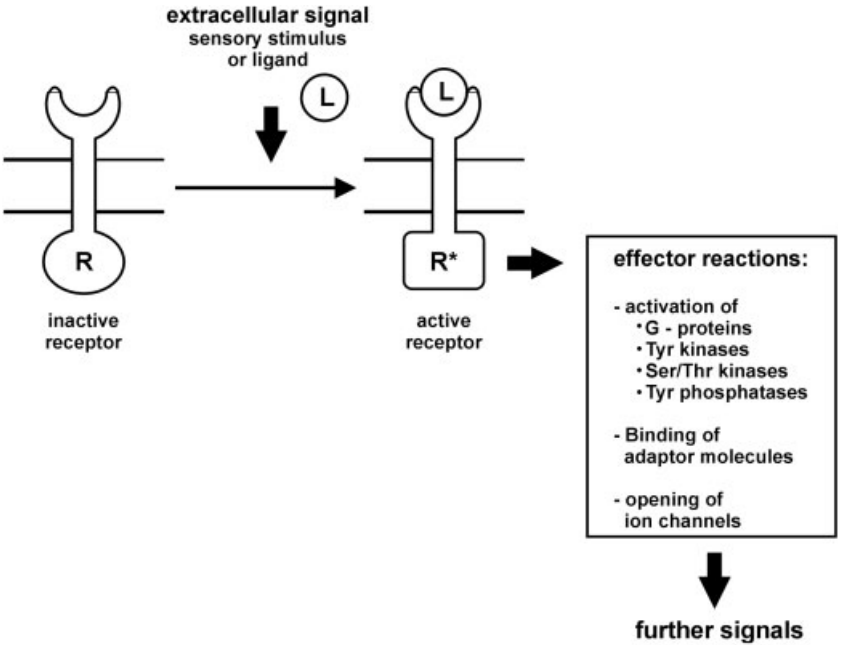


Fig. 5.5 General functions of transmembrane receptors. Extracellular signals convert the transmembrane receptor from the inactive form R to the active form R*. The activated receptor transmits the signal to effector proteins next in the reaction sequence. Important effector reactions are the activation of heterotrimeric G-proteins, of protein tyrosine kinases and of protein tyrosine phosphatases. The tyrosine kinases and tyrosine phosphatases may be an intrinsic part of the receptor or they may be associated with the receptor. The activated receptor may also include adaptor proteins in the signaling pathway or it may induce opening of ion channels.

Starting from the activated receptor, a large number of reactions can be set in motion (Fig. 5.5). One main route of signal transmission takes place by activation of G proteins, another via activation of tyrosine-specific protein kinases, and a further route is via activation of ion channels. In the further course of G protein-mediated signal transmission, secondary diffusible signals are often formed: the "second messenger" molecules (see Chapters 3 and 6). These function as effectors and activate further enzyme systems in the sequence, especially protein kinases.

The activated receptor can also associate with adaptor molecules, which serve as coupling elements for further signal proteins.

5.2.4 Regulation of Receptor Activity

A physiologically important aspect of signal transmission via transmembrane receptors is its regulation. The cell has various mechanisms available, with the help of which the number and activity of transmembrane receptors can be regulated. The aim of

regulation is, for example, to weaken signal transmission via the receptor during conditions of long-lasting hormonal stimulation. Furthermore, signal transduction by transmembrane receptors may be modulated via crosstalk with other signaling pathways. The structural elements involved in regulation of receptor activity are generally located in the cytosolic domain. These are, above all, protein sequences that permit phosphorylation of the receptor by protein kinases. Phosphorylation at Ser/Thr or Tyr residues of the cytosolic domain may lead to inactivation or activation of the receptor and thus weaken or strengthen signal transmission. In this way, Ser/Thr-phosphorylation is used in the process of internalization of receptors in order to remove the receptor from circulation after it has been activated (see Section 5.3.4). The protein kinases involved are often part of other signaling pathways and can link the activity of the transmembrane receptors to other signaling networks.

Targeted degradation of transmembrane receptors is another means of regulating receptor activity. Signals for ubiquitination and subsequent degradation in the proteasome have been identified in the cytosolic domain of transmembrane receptors.

5.3 G Protein-Coupled Receptors

Of the transmembrane receptors that receive signals and conduct them into the cell interior, the G protein-coupled receptors form the largest single family. About 5 % of the genome of the worm *Caenorhabditis elegans* is occupied by genes encoding G protein-coupled receptors. In vertebrates, more than 1000 different G protein-coupled receptors are found that may be activated by extracellular ligands or sensory signals. The ligands include biogenic amines, such as adrenaline and noradrenaline, histamine, serotonin, lipid derivatives, nucleotides, retinal derivatives, peptides such as bradykinin and large glycoproteins such as luteinizing hormone, and parathormone (see also Table 3.1).

Activation of receptors can also be mediated by proteolytic cleavage of the extracellular domain of the receptor by proteases like thrombin. For these *protease-activated receptors*, a proteolytically produced peptide functions as the activating ligand.

In addition, physical stimuli such as light signals are registered and converted into intracellular signals by G protein-coupled receptors; they are also involved in perception of taste and smell.

Ligand binding or reception of a physical signal is linked to activation of the G protein-coupled receptor. As a consequence, the receptor undergoes a conformational change that is transmitted to the inner side of the membrane, whereby the next sequential member of the signal chain is activated. This conducts the signal further via other reaction pathways (see Fig. 5.16). A characteristic structural feature of the G protein-coupled receptors is the presence of 7 transmembrane helices. For the vast majority of 7-helix transmembrane receptors the next downstream located signaling protein is a heterotrimeric G protein. This is the origin of the name of these receptors. However, alternative routes for signal transmission from activated G protein-coupled

Tab. 5.1 Classification of the heterotrimeric G proteins according to the α-subunits.

Subunit	tissue	Examples of receptors	Effector protein, function
G_s			
α_S	ubiquitous	βAR, glucagon receptor	adenylyl cyclase ↑ Ca^{2+} channels ↑
α_{olf}	nasal epithelium	olfactory receptor	adenylyl-cyclase ↑
G_i			
α_{i1}, α_{i2}, α_{i3}	mostly ubiquitous	α_2 adrenergic receptor α_2 adrenergic receptor	K^+ channels ↑ Ca^{2+} channels ↓ K^{2+}-channels ↑ Ca^{2+}-channels
α_{t1} transducin	retina	rhodopsin	eGMP specific phosphodiesterase ↑
α_g	taste buds		
α_2	brain		adenylyl cyclase ↓
G_q			
α_q $_{11}$, α_{14}, α_{15}, α_{16}	ubiquitous	α_1 adrenergic receptor	phospholipase Cβ ↑
G_{12}			
α_{12}, α_{15}	ubiquitous		

receptors have been discovered. One of these leads to the MAPK cascade (see below: Chapter 10), resulting in the activation of transcription factors. Another way links signaling by G protein-coupled receptors to signaling by growth factor receptors. Stimulation of G protein-coupled receptors by extracellular ligands like lysophosphatidic acid results in a transactivation of the receptor for epidermal growth factor, EGFR (review: Gschwind et al., 2001). The mechanistic basis for this cross-communication between two different signaling pathways is not well understood and involves proteolytic processing of the extracellular ligand by a metalloprotease.

5.3.1 Structure of G Protein-Coupled Receptors

The key structural elements of the G protein-coupled receptors are seven sequence segments, each made up of 20–30 amino acids, that form a transmembrane domain, and span the membrane in the form of α-helices. The transmembrane elements are linked by loops of various sizes on the outer and inner side. The highest sequence homology of G protein-coupled receptors is found in the transmembrane elements, whilst the hydrophilic loop regions show stronger divergence between different receptors.

Examplary for the G-protein-coupled receptors, Fig. 5.6 shows the two dimensional model of bovine rhodopsin. The three-dimensional structure of rhodospsin has been directly visualized high resolution X-ray analysis (Palczewski et al., 2000; see Fig 5.3)

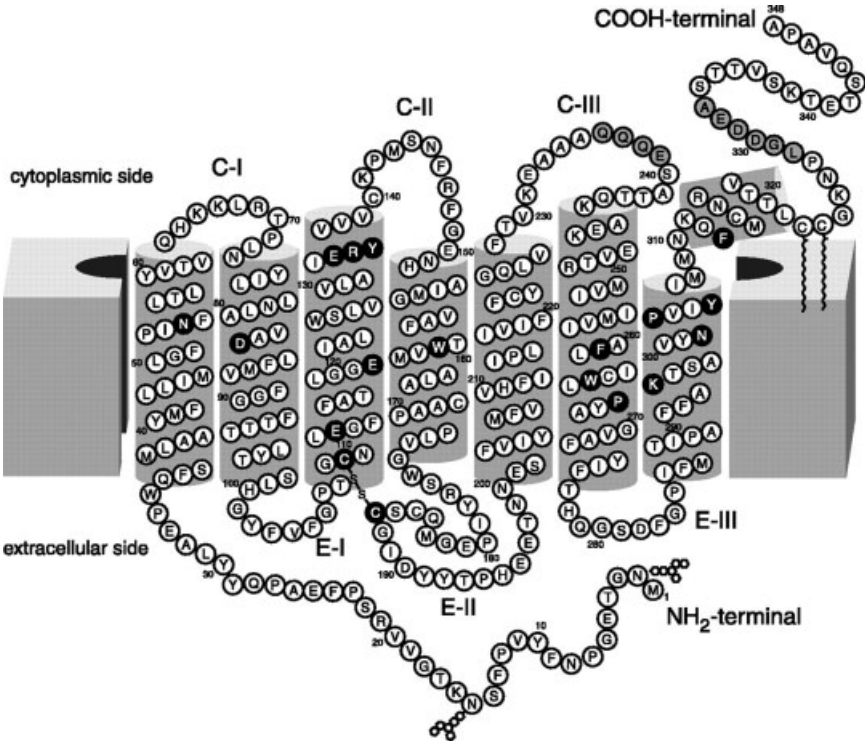


Fig. 5.6 Two dimensional model of rhodopsin. The extracellular (intradiscal) and intracellular regions of rhodopsin each consist of three interhelical loops (given the prefixes E(extracellular)-I to E-III or C(cytoplasmic)-I to C-III. A conserved disulfide bridge is found on the extracellular side linking EII with E-III. On the intracellular side, a short helix runs parallel to the membrane surface. In the native protein, the C-terminus carries two palmitoylated Cys-residues which function as membrane anchors causing formation of a putative fourth intracellular loop.

that shows a bundle of 7 transmembrane helices, as predicted from a multitude of biochemical and biophysical studies. This structure provides a frame upon which the 3D-structure of the huge family of 7-helix transmembrane receptors can be modeled. Sequence comparisons, biochemical and biophysical data indicate that the transmembrane bundle structure is conserved among GPCRs and possibly within the entire GPCR family. By contrast, the loops and termini are more divergent in amino acid sequence and possibly in three-dimensional structure.

The rhodopsin structure shows a bundle of 7 transmembrane helices that vary in length from 20 to 33 residues. Another helix, helix VIII, is found on the intracellular side and runs parallel to the membrane. In the extracellular region the structure is highly organized and includes a conserved disulfide bridge. The chromophore 11-*cis* retinal, which mediates the light activation, is covalently bound in the interior of the bundle. It keeps the structure in the ground state. Light absorption excites the chromophore and activates the rhodopsin, probably by reorienting the helices with-

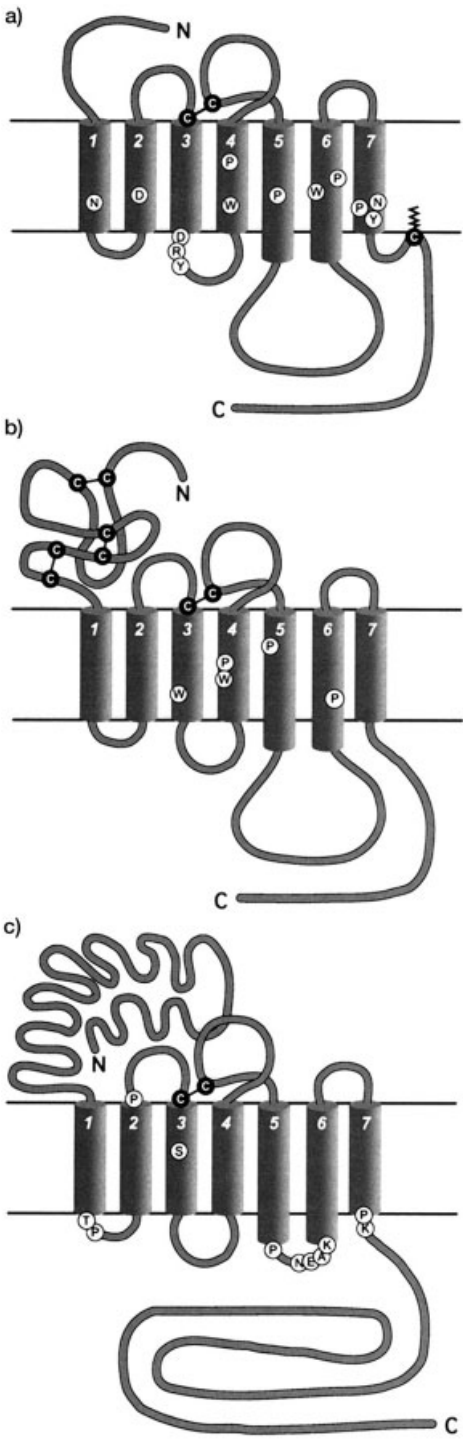


Fig. 5.7 Classification of GPCRs. The G protein-coupled receptors can be divided into three major subfamilies (see Gether, 2001)
Family A receptors are characterized by a series of highly conserved key residues (black letter in white circles). In most family A receptors, a disulfide bridge is connecting the E-II and E-III loops. In addition, a majority of the receptors have a palmitoylated cysteine in the cytoplasmic C-terminus.
Ligands include the biogenic amines (adrenaline, serotonine, doapmine, histamine), neuropetide Y, adenosine, chemokines and melatonine, among others.
Family B receptors are characterized by a long extracellular N-terminus containing a series of cysteine residues presumably forming a network of disulfide bridges. Representative members of the family B receptors include calcitonine receptor, glucagon receptor and parat hormone receptors.
Family C receptors are characterized by a very long N-terminus forming the extracellular ligand binding site. There is only one putative disulfide bridge and the third cytoplasmic loop is very small.
The taste receptors, the metabotropic glutamate receptors, the γ-aminobutyric acid (GABA) receptors and Ca^{2+}-receptors belong to this class, among others.

in the bundle. Activation of the downstream effector protein, the G protein, is mediated by the intracellular part of rhodopsin. This region includes three loops, helix VIII and the C-terminus and forms an extended surface for the interaction with the downstream effector, the G protein. Details of the activation process are however still unknown.

The only structural feature common to all G protein-coupled receptors is the presence of the seven transmembrane helices connected by alternating extracellular and intracellular loops, with the amino terminus located on the extracellular side and the carboxy terminus on the intracellular side. Apart from that feature, the overall sequence homology among the G protein-coupled receptors is low. Significant sequence homology is found, however, within three subfamilies, designated family A, B and C receptors (review: Gether, 2000). The classification is based on the size of the extracellular loops, the presence of key residues and the formation of disulfide bonds (Fig. 5.7). Family A includes the rhodopsin/β-adrenergic receptor, family B includes calcitonin receptors and family C includes receptors for γ-amino butyric acid, Ca^{2+} and glutamate.

The 7-helix transmembrane receptors are often *glycoproteins*. Glycosylation sites are located in the extracellular region, e.g., in the form of the consensus sequence Asn-X-Ser/Thr for an N-linked glycosylation.

The extracellular loops contain frequently *conserved Cys residues*. It is assumed that these stabilize the conformation of the extracellular domain, via disulfide bridges.

Post-translational modification in the form of *palmitoylation* of the cytosolic domain (see Section 3.7.2) has been demonstrated for the α- and β-adrenergic receptor. The palmitoylation takes place on a Cys residue localized at the C-terminus on the membrane inner side. The modification may possibly serve to anchor the C-terminus in the membrane.

5.3.2 Ligand Binding

The area of ligand binding has been particularly well defined for the receptors of classical "small ligands" (adrenaline, noradrenaline, dopamine, serotonine, histamine). Targeted mutagenesis, biochemical, biophysical and pharmacological investigations have shown that these ligands are bound in a binding crevice formed by the transmembrane helices. In agreement with this model, it has been shown that the extracellular and intracellular sequence portions of the receptors are not needed for ligand binding in these cases. Rhodopsin is a unique case, since the retinal ligand is covalently attached by Schiff-base linkage to a Lys residue of transmembrane helix VII. In that case too, the binding site is deeply buried in the interior of the transmembrane segment (see Fig 5.3).

For the receptors that have peptides or proteins as ligands, structural portions of the extracellular domain, in addition to areas of the transmembrane domain, are involved in ligand binding. The receptors of subfamily C that contain a large extracellular N-terminal domain bind their ligands (e.g., glutamate, γ-amino butyric acid) in this region.

5.3.3
Mechanism of Signal Transmission

The mechanism by which the activated receptor talks to the G protein is still a matter of speculation. Generally, the switch function of the receptor is considered in terms of allosteric conformational changes within the 7-helix membrane bundle (review: Meng and Bourne, 2001; Okada et al., 2001). According to this representation, changes in the structure of the transmembrane bundle are passed on to the cytoplasmic loops of the receptor. As a consequence, a high-affinity surface is created for binding of the heterotrimeric G protein (see Section 5.5.4). The heterotrimeric G protein, which exists as the inactive GDP form, now binds via its α- and possibly γ-subunit to the activated receptor and is activated itself. An exchange of GDP for GTP takes place, and the $\beta\gamma$-subunit of the G protein dissociates (see Section 5.5.3). Once the G protein is activated, it frees itself from the complex with the receptor, which either returns to its inactive ground state or activates further G proteins.

5.3.4
Switching Off and Desensitization of 7-Helix Transmembrane Receptors

A phenomenon often seen in transmembrane receptors in general, and in G protein-coupled receptors in particular, is desensitization (Fig. 5.8). *Desensitization* means a weakening of the signal transmission under conditions of long-lasting stimulation by hormones, neurotransmitters or sensory signals. Despite the persistent effect of extracellular stimuli, the signal is no longer passed into the cell interior, or only in a weakened form, during desensitizing conditions. This is a mechanism with which both short-term and long-term regulation of receptor activity is possible.

The best investigated is the desensitization of the adrenaline receptor type β2 and of rhodopsin. Rhodopsin has the function of a light receptor in the process of vision. It

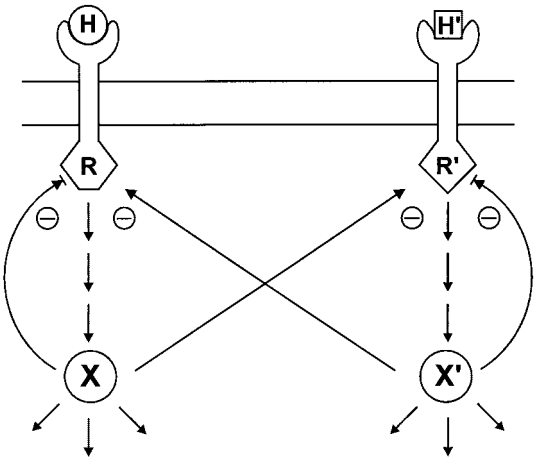


Fig. 5.8 General principle of densensitization of G-protein-coupled receptors. Desensitization of a hormone-bound receptors can take place by two principle routes, schematically represented in the figure. A suppressing influence may be exerted on the receptor system via proteins (X) of a signal chain, triggering inhibition of the signal chain. Receptor systems may also mutually influence one another in that a signal protein X formed in one signal chain mediates the densensitization of another receptor system R*, and vice versa.

receives light signals and conducts them to the relevant G protein, transducin. The *key reaction* in desensitization of both systems is the phosphorylation of the receptor at the cytoplasmic side by specific protein kinases.

Furthermore, Ser/Thr phosphorylation can be used as a switch for coupling a given receptor to different G_α subunits. Protein kinase A-mediated phosphorylation of the β-adrenergic receptor has been shown to switch coupling of the receptor from G_s to G_i and initiate a new set of signalling events (Daaka et al., 1997).

In addition, phosphorylation-mediated binding of arrestins to the receptor (see below) serves as a means to couple G protein-coupled receptor to another signaling pathway, the MAPK cascade.

Two classes of protein kinases are mainly involved in the phosphorylation and desensitization (review: Freedman and Lefkowitz, 1996):

- Phosphorylation by cAMP-dependent protein kinases (Fig. 5.8)
  Phosphorylation of the cytoplasmic domain of 7-helix transmembrane receptors can take place via *cAMP-dependent protein kinases* (*protein kinase A*) or via *protein kinase C* (Chapter 7) (Fig. 5.9). This is a feedback mechanism. The hormonal activation of the receptor leads, via G proteins and adenylyl cyclase/cAMP, to activation of protein kinases of type A (see Sections 5.6.1 and 6.1, and Chapter 7). The activated protein kinases phosphorylate the receptor in the region of the cytoplasmic domain on Ser/Thr residues. Regulation via adenylyl cyclase/cAMP/proteinkinase A is an example of a heterologous desensitization, since adenylyl cyclase can be activated by a variety of signals originating from different signaling pathways (see Section 5.6.1).

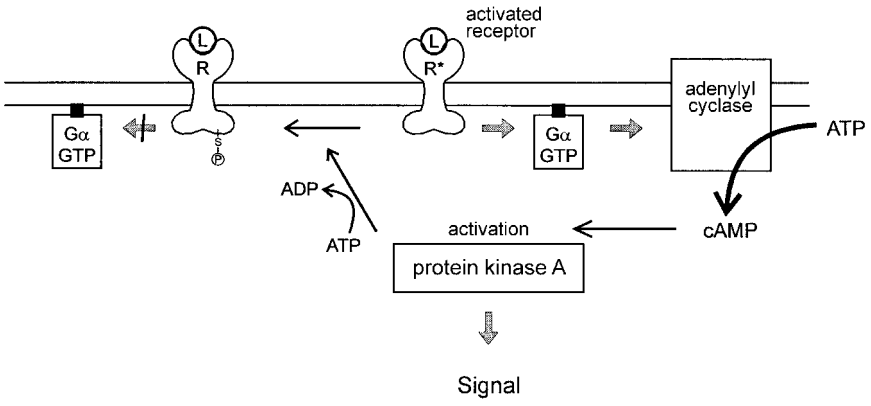


Fig. 5.9 Desensitization of G-protein-coupled receptors via cAMP-dependent protein kinases. Starting from an activated receptor R*, the signal is transmitted via the G_α-subunit of the G-protein to adenylyl cyclase. The latter is activated and forms cAMP. This activates a protein kinase of type A that passes the signal in the form of a Ser/Thr-specific protein phosphorylation to substrate proteins. One of the substrates is also the receptor that is phosphorylated in the region of the cytoplasmic domain by the activated protein kinase A. The ligand-bound receptor is preferentially phosphorylated. As a consequence of phosphorylation, activation of further G-proteins by the receptor is suppressed.

- Phosphorylation via G protein coupled receptor protein kinases (GRK)
  The major mechanism for the homologous desensitization of agonist-bound 7-helix transmembrane receptors consists of a two-step process in which the agonist-bound receptor is phosphorylated by a GRK and then binds an arrestin protein which interrupts signaling to the G protein. Well-characterized GRKs (review: Pitcher et al., 1998) are those for rhodopsin, rhodopsin kinase and the β2 adrenaline receptor, the β-adrenergic receptor kinase (βARK). The GRKs are protein kinases that are

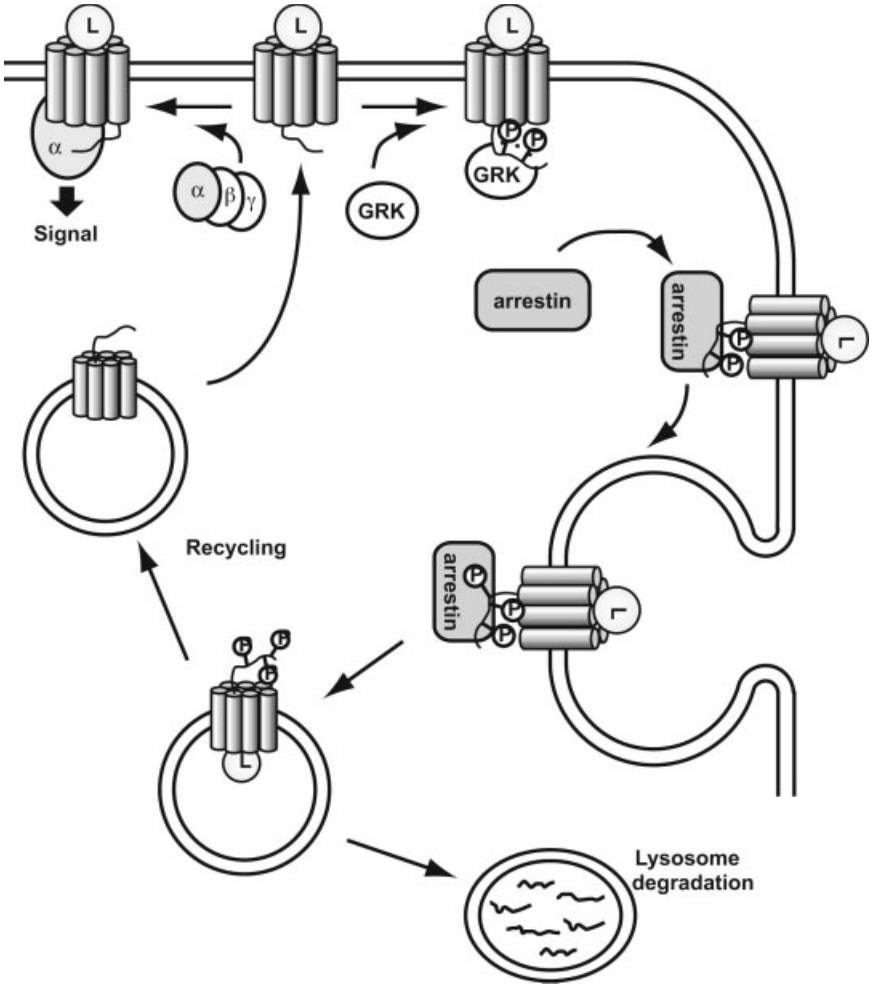


Fig. 5.10 Receptor desensitization: phosphorylation, arrestin binding and internalization. The activated, agonist-bound receptor is phosphorylated on the cytoplasmic region by a G protein-coupled receptor protein kinase (GRK). The phosphate residues serve as attachment sites for β-arrestin which has protein kinases of the MAPK cascade associated. This serves as a trigger for internalization of the receptor to endosomes. The receptor may now be dephosphorylated and transported back to the cell membrane (not shown in the figure).

not regulated via second messengers. Phosphorylation of the receptor by GRKs takes place in the C-terminal region and/or in the third cytoplasmic domain. Only the activated, i.e. occupied by an agonist, receptor is phosphorylated. Activation of the GRK is linked to its transport from the cytosol to the inner side of the membrane, where it phosphorylates the receptor. During translocation of most GRKs to the membrane-localized receptor, the $\beta\gamma$-subunit of the G protein as well as binding of phosphatidyl inositol triphosphates (see Section 6.6.2) play an important role (see Fig. 5.9 and Section 5.5.7). When an agonist binds to a G protein-coupled receptor, it causes the receptor to associate a heterotrimeric G protein, leading to dissociation into its α and $\beta\gamma$ dimer subunits. The $\beta\gamma$ dimer, which carries a prenyl membrane anchor, binds to the GRK and thereby mediates its membrane association. In a cooperative way, binding of phosphatidyl-inositol-messengers (see Chapter 6) to the PH-domain (see Section 8.2.4) of GRKs enhances binding of the GRK to the membrane.

Receptor phosphorylation by GRKs triggers several reactions (Fig. 5.10):

Binding of Arrestin

Phosphorylation of the receptor by GRKs leads to creation of a high-affinity binding site for proteins known as *arrestins*. The family of arrestins includes the visual arrestins that are specific for rhodopsin and the β-arrestins that bind, e.g., to the β-adrenergic receptor. This class of β-arrestins is most important for the desensitization of the activated receptor. Binding of β-arrestins has several consequences:

- The phosphorylated receptor is decoupled from the interaction with the heterotrimeric G protein next in the sequence so that signal transmission is suppressed. Arrestin binding serves, e.g., to rapidly weaken signal transmission during the vision process, during conditions of long lasting light stimulus.
- Following β-arrestin binding, the phosphorylated receptor is translocated into the cell interior. In this process, β-arrestin serves as an adaptor that targets the receptor for internalization via endocytosis. In addition to receptor binding, the β-arrestins bind to clathrins and other proteins that form the clathrin-coated vesicle machinery. The receptor is thereby internalized in the membrane-associated form, dephosphorylated, and then transported back to the cell membrane. The translocation into the cell interior serves, in particular, to weaken signal transmission during conditions of long-lasting hormonal stimulation.
- For some receptors, arrestin binding serves to activate a protein kinase cascade, the MAPK cascade (review: McDonald and Lefkowitz, 2001. In that signaling mode, arrestins function as scaffolding proteins that help to organize the three protein kinases of the MAPK pathway into a cascade of sequentially acting protein kinases delivering a signal to the level of, e.g., transcription factors (see Chapter 10). In view of this observation, arrestins appear to function not only as "signal terminators" but rather also as activators of another signaling pathway, that of the MAPKs, providing another example of a crosstalk between different signaling pathways (Fig. 5.11).

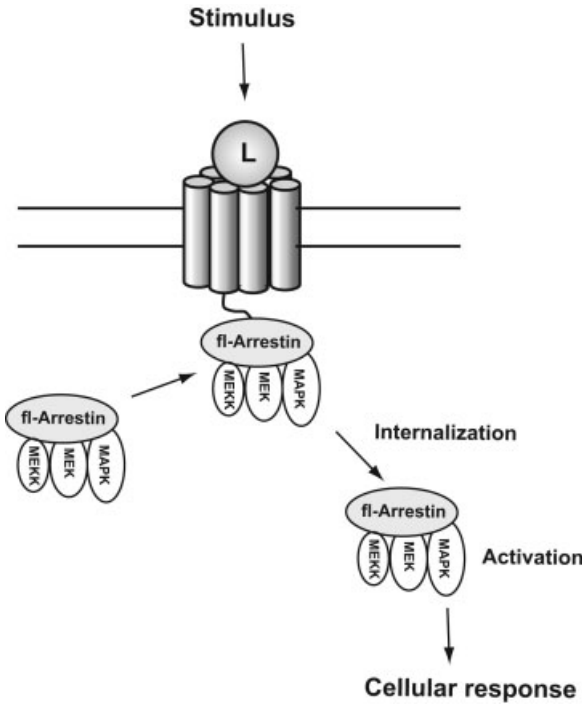


Fig. 5.11 Activation of the MAPK cascade via β-arrestin and G protein-coupled receptors. A complex is formed between β-arrestin and the various components of a MAPK module (see Chapter 10) in the cytosol. This multiprotein complex translocates to the plasma membrane bound receptor following ligand binding. β-arrestin functions as a scaffold for the MAPK module and promotes internalization of the whole complex. This leads to generation of active MAPK and stimulation of transcription. MAPK: mitogen activated protein kinase, MEK: MAPK/ERK kinase, MEKK: MEK kinase.

Signal Switching

In another reaction, Ser/Thr phosphorylation of 7-helix transmembrane receptors can be used as a switch for coupling a given receptor to different G_α subunits, the downstream effector protein. Protein kinase A-mediated phosphorylation of the β-adrenergic receptor has been shown to switch coupling of the receptor from G_s to G_i and trigger a new set of downstream signaling reactions (Daaka et al., 1997).

5.3.5 Dimerization of GPCRs

Although the GPCRs were generally believed to function as monomeric entities, there is now increasing evidence that G protein-coupled receptors may form functional dimers *in vivo*. Biochemical and biophysical studies suggest that, e.g., the β2 adrenergic receptor exists as a constitutive dimer in the cell (Angers et al., 2002). Dimerization has been shown to alter the ligand-binding, signaling and trafficking properties of G protein-coupled receptors. In addition to homodimers, the formation of heterodimers with related members of the same subfamily has been also reported. The structural, functional and mechanistic consequences of the formation of the oligomeric receptor complexes remain to be elucidated.

5.4
Regulatory GTPases

The heterotrimeric G proteins, the major effector proteins of the 7-helix transmembrane receptors, belong to the large family of regulatory GTPases; these bind GTP and hydrolyze it, thereby functioning as a switch in central cellular processes. The family of regulatory GTPases is also called the GTPase superfamily.

5.4.1
The GTPase Superfamily: General Functions and Mechanism

Proteins of the GTPase superfamily are found in all plant, bacterial and animal systems. The following examples illustrate the central functions of the regulatory GTPases in the cell.

Regulatory GTPases are involved in
- protein biosynthesis on ribosomes
- signal transduction at membranes
- visual perception
- sense of smell and taste
- control of differentiation and cell division
- translocation of proteins through membranes
- transport of vesicles in the cell.

The members of the GTPase superfamily show an extensively conserved reaction mechanism. A common trait is a switching function that enables a reaction chain to be switched on or off.

The Switch Function of the GTPases

The regulatory GTPases are involved in signaling chains by functioning as a switch. Incoming signals are received by the GTPases and are passed on to downstream components of the signaling chain. The switch function is based on a cyclical, unidirectional transition between an *active, GTP-bound form* and an *inactive, GDP-bound form* (Fig. 5.12). The binding of GTP brings about the transition into the active form. Hydrolysis of the bound GTP by an intrinsic GTPase activity converts the protein into the inactive, GDP-bound form. In both inactive and active forms, the proteins of the GTPase superfamily possess a specific affinity to other proteins employed earlier or later in the reaction chain. In the active GTP form, the signal is passed on to the effector molecule next in sequence. Conversely, if the GTPase is in the inactive, GDP-bound form, the reaction chain is disrupted. To reactivate the reaction chain, the GDP must be exchanged for GTP. This step is linked to the receipt of an incoming signal that triggers the exchange of GDP for GTP. The dissociation of GDP from the inactive complex is a process that generally occurs by interaction with upstream protein components of the signaling chain.

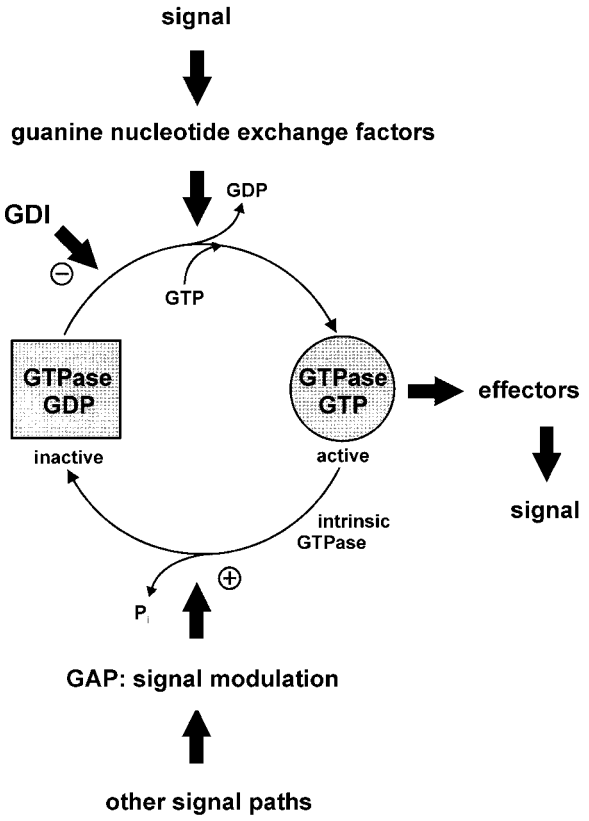


Fig. 5.12 The switch function of the regulatory GTPases. The GTP form of the regulatory GTPases represents the "switched on" form of the GTPase, the GDP form, in contrast, the "switched off" form. The switch function of the regulatory GTPases may be controlled by guanine nucleotide exchange factors, by GTPase activating proteins (GAPs) and by G-nucleotide dissociation inhibitors (GDIs). The regulatory GTPases run through a GTPase cycle which signals flow into via GEFs and are conducted further in the form of the GTPase-GTP complex to effector molecules further down the sequence. Hydrolysis of the bound GTP ends the activated state. The rate of GTP hydrolysis is either intrinsically determined or may be accelerated via GAPs.

The GTPase Cycle

The GTPases bring about the transition between the active and inactive states in a *cyclic process* that can only run in one direction because of the irreversible hydrolysis of GTP. At least three different GTPase states can be differentiated in the GTPase cycle: the active GTP form, the inactive GDP form and an "empty" form of the GTPase, which is generally a short-lived state.

The switch function of the GTPase is based on the specific ability of the different functional states of the GTPase to interact with the proteins that precede and follow in the signal chain. Thereby the regulatory GTPases receive and transmit signals functioning as a switch in the signaling chain. An incoming signal is received by the GDP-bound, inactive state, which is thereupon converted into the active GTP state. From this state the signal is passed on to the next effector protein in place. The activated state is switched off by hydrolysis of GTP, whereby the GTPase returns to its inactive ground state. A particular GTPase is characterized by the proteins with which the active and inactive forms interact.

Modulation and Regulation of the Switch Function

How effectively a signal can be transmitted by a GTPase depends on the relationship of the concentration of the active GTP form to the inactive GDP form. This, in turn, is determined by the relationship of the rate constant for the dissociation of GDP, $k_{diss,GDP}$ to the rate constant of GTP hydrolysis, $k_{cat,GTP}$:

$$\frac{\text{GTPAse} \cdot \text{GTP}}{\text{GTPAse} \cdot \text{GDP}} = \frac{k_{\text{diss. GDP}}}{k_{\text{cat. GTP}}}$$

This relationship is valid if it can be assumed that the GTP concentration is not limited and that GTP binds very rapidly to the "empty" form of the GTPase. A special characteristic of the regulatory GTPases is that both rate constants may be regulated by specific proteins. The proportion of GTPase that exists in the active form can be altered by at least three processes:

1. Acceleration of the dissociation of GDP increases the proportion of the active form. The rate of dissociation of GDP may be increased by specific proteins. These proteins are known as *guanine nucleotide exchange factors* (GEF). For the heterotrimeric G proteins, the agonist-bound, activated receptor is the exchange factor.
2. Dissociation of GDP may be inhibited by specific proteins known as *guanine nucleotide dissociation inhibitors* (GDI). Proteins with this function are found in members of the superfamily of Ras proteins (see Chapter 9). The GDIs have the function, above all, to provide a cytosolic pool of inactive, GDP-bound proteins (see Section 9.1).
3. Increasing of the rate of GTP hydrolysis by *GTPase activating proteins* (GAP) reduces the lifetime of the active, GTP-bound state. The GAP protein class is an important instrument for control of the rate of signal transmission. Activation of the GAPs leads to strengthening or weakening of signal transmission. Often, the activity of the GAPs is regulated by other signaling pathways. Thus, a regulatory influence on signal transmission via G proteins can be achieved from another signaling pathways.

The various GTPases may differ to a large extent in the rate of GTP hydrolysis and thereby in the influence of GTPase-activating proteins. The Ras protein and the $G_{\alpha t}$, known as transducin, involved in the process of vision, are cited as examples.

The Ras protein has low intrinsic GTPase activity. This may be increased ca. 10^5-fold by the corresponding GTPase-activating protein (see also Chapter 9). In comparison, the intrinsic rate of GTP hydrolysis of transducin is ca. 100-fold higher than that of the Ras protein. The effector molecule downstream of transducin, the cGMP phosphodiesterase, is involved in GTPase stimulation in this case. The γ-subunit of the phosphodiesterase functions in cooperation with a member of the RGS protein famliy (see Section 5.5.7), namely RGS9, as the GAP here and stimulates GTPase activity of the transducin by about two orders of magnitude.

5.4.2
Inhibition of GTPases by GTP Analogs

Nonhydrolyzable GTP analogs are an indispensable tool in the identification and structural and functional characterization of GTPases. The GTP analogs shown in Fig. 5.13, GTPγS, β,γ-methylene GTP and β,γ-imino GTP, are either not hydrolyzed by GTPases or only very slowly. Addition of these analogs fixes the G protein in the active form; it is permanently "switched on". For cellular signal transduction, this means permanent activation of the signal transmission pathway. In many cases, a role of G proteins in a signal chain was inferred from the observation that nonhydrolyzable GTP analogs bring about a lasting activation of signal transmission. The GTP analogs were equally important for structural determination of the activated form of GTPases. Formation of a stable complex between the nonhydrolyzable GTP analog and different GTPases has enabled crystallization of the complex in its activated form.

5.4.3
The G-domain as Common Structural Element of the GTPases

A common property of the GTPases is the enzymatic activity of GTP hydrolysis. GTP binding and hydrolysis take place in a domain of the GTPases known as the *G-domain*.

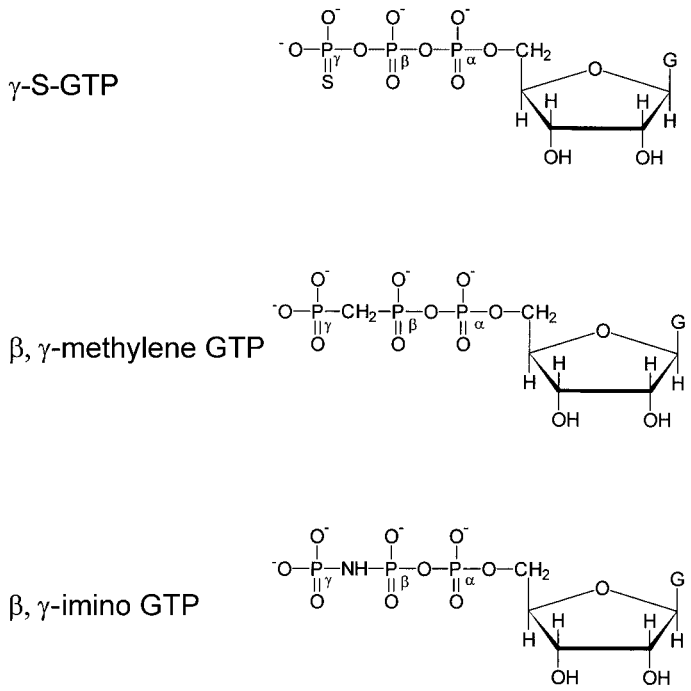


Fig. 5.13 Examples of non-hydrolysable GTP analogs.

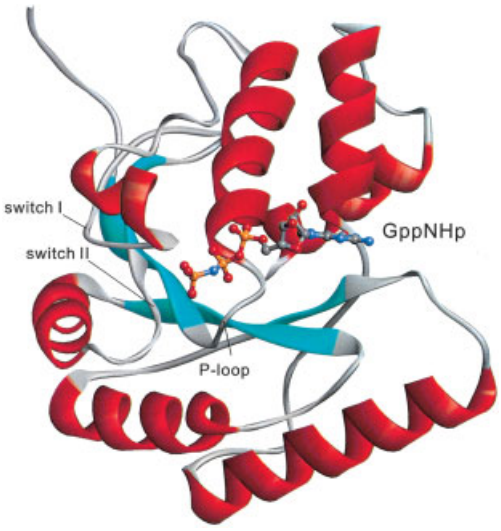


Fig. 5.14 Structure of the G-domain of the elongation factor EF-Tu from *T. thermophilus* with bound GppNHp, according to Berchthold et al., (1993). The non-hydrolysable analog GppNHp, the P loop and the switch regions I and II are shown, which play an important role in transition from the inactive GDP form to the active GTP form.

The G-domain is found in all GTPases. Figure 5.14 shows the G-domain of the bacterial elongation factor EF-Tu. In all GTPase structures known at present, the G-domain has very similar architecture and very similar means of binding the guaninine nucleotide. The sequence element $GX_4GK(S/T)$ is a consensus sequence for guanine nucleotide binding; this sequence is involved in binding the β- and γ-phosphate of GTP and GDP and is also known as the P-loop. Other consensus sequences, such as RX_2T and DX_2G, are involved in both binding the γ-phosphate and in the GTPase reaction (X = any amino acid). A further consensus sequence (N/T)(K/Q)XD and SA interacts with the guanosine.

5.4.4 The Different GTPase Families

The superfamily of GTPases, with over 100 members, is subdivided according to sequence homologies, molecular weight and subunit structure into further (super)families. These are the families of the *heterotrimeric G proteins*, the *Ras superfamily of small GTPases* and the *family of initiation and elongation factors* (Fig. 5.15).

The heterotrimeric G proteins are built of up three subunits, with the GTPase activity localized on the largest subunit (see Section 5.5). The members of the Ras family of GTPases, in contrast, are monomeric proteins with a molecular weight of ca. 20 kDa (see Chapter 9).

A further functionally diverse class is made up of the proteins involved in protein biosynthesis and membrane transport. GTPases with functions in protein biosynthesis include the elongation factors, termination factors and peptide translocation factors. These are mostly monomeric proteins with molecular weights of 40 – 50 kDa.

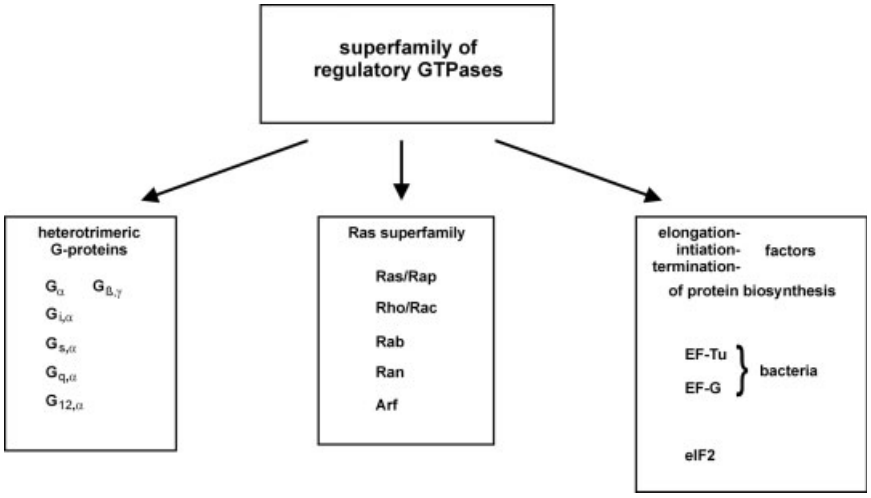


Fig. 5.15 The GTPase superfamily.

GTPases of this class are also found in protein complexes such as the "signal recognition particle" (SRP) and the corresponding receptor. Both protein complexes are needed during ribosomal protein biosynthesis, for transport of newly synthesized proteins through the endoplasmic reticulum.

5.5 The Heterotrimeric G Proteins

The heterotrimeric G proteins are the specific reaction partners in signal transmission via 7-helix transmembrane receptors, which is why these receptors are also known as G protein-coupled receptors. From the G protein, the signal is then passed on to the effector protein next in the sequence.

A common structural feature of the G proteins is their construction from three subunits (Fig. 5.16), a large α-subunit of 39–46 kDa, a β-subunit of 36 kDa and a γ-subunit of 8 kDa. The α-subunit has a binding site for GTP or GDP and carries the GTPase activity. The β- and γ-subunits exist as a tightly associated complex and are active in this form. All three subunits show great diversity, so that at least 20 different genes for α-subunits, 5 for β-subunits and 12 for γ-subunits are known in mammals. Some G protein are ubiquitous, whereas others only occur in specialized tissue.

Specificity of the switch function is mostly determined by the α-subunit: the α-subunit carries out the specific interaction with the receptors preceding it in the signal chain and with the subsequent effector molecules. The $\beta\gamma$-complex is also involved in signal transmission to the effector proteins.

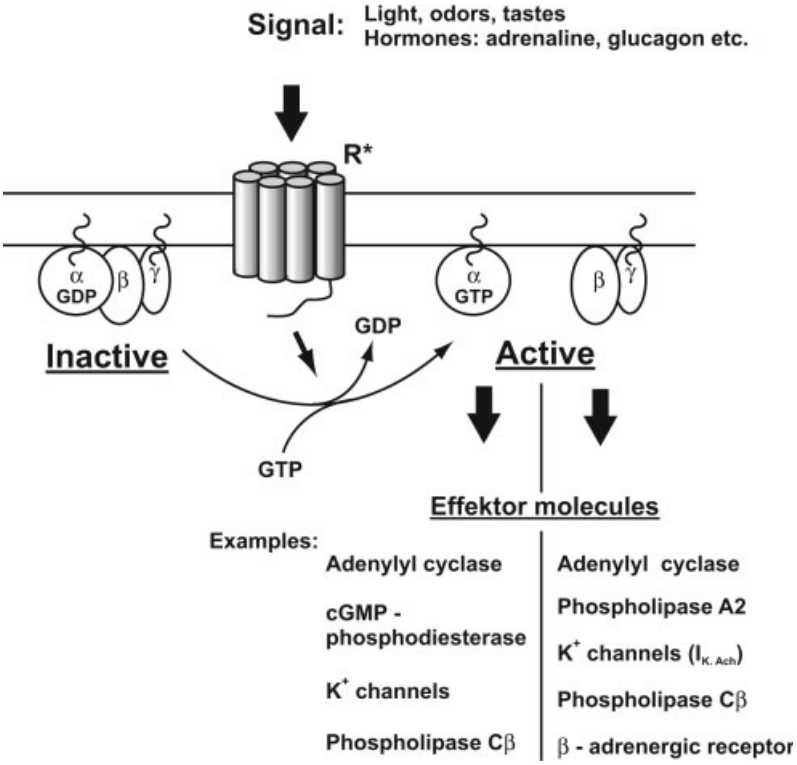


Fig. 5.16 Structure and activation of the heterotrimeric G-proteins. Reception of a signal by the receptor activates the G-protein, which leads to exchange of bound GDP for GTP at the α-subunit and to dissociation of the $\beta\gamma$-complex. Further transmission of the signal may take place via G_α-GTP or via the $\beta\gamma$-complex, which interact with corresponding effector molecules. The α- and γ-subunits are associated with the cell membrane via lipid anchors. Signal reception and signal transmission of the heterotrimeric G-proteins take place in close association with the cell membrane. This point is only partially shown in the figure.

5.5.1 Classification of the Heterotrimeric G Proteins

Most functions of signal transmission by G proteins are realized by the α-subunit. Since different G proteins interact with very different partners, there are significant differences in the structure of the α-subunits. Because of the common GTPase domain and the common interaction with the $\beta\gamma$-subunits, however, there are also considerable sequence homologies. Based on comparison of the amino acid sequences, the Gα-proteins are divided into *four families, the G_s, G_i, G_q and G_{12} families* (review: Downes and Gautam, 1999). These families are summarized in Table 5.1, together with representative members and their characteristic properties. Classification based on homology of amino acid sequences, however, does not give any information about the functional properties of the different α-subunits.

G_s Subfamily

A characteristic of α-subunits of the G_s subfamily is that they are inhibited by cholera toxin (see Section 5.5.2). The members of the G_s subfamily are activated by hormone receptors, by odor receptors and by taste receptors. G_s-proteins mediate, e.g., signal transmission by type β adrenaline receptors and that by glucagon receptors. During perception of taste, the taste receptors are activated, and these then pass the signal on via the olfactory G protein G_{olf}. Perception of "sweet" taste is also mediated via a G_s-protein. Transmission of the signal further involves an adenylyl cyclase in all cases, the activity of which is stimulated by the G_s-proteins.

G_i Subfamily

The first members of the G_i subfamily to be discovered displayed an inhibitory effect on adenylyl cyclase, hence the name G_i, for inhibitory G proteins. Further members of the G_i subfamily have phospholipase C as the corresponding effector molecule. Signal transmission via phospholipase C flows into the inositol triphosphate and diacylglycerol pathways (see Chapter 6).

The G_t- and G_g -proteins are also classed as G_i-proteins, based on sequence homologies. The G_t- and G_g -proteins are involved in transmitting sensory signals. Signal transmission in the vision process is mediated via G proteins known as *transducins* (G_t). The G_t-proteins are activated by the photoreceptor rhodopsin and are located in the rods and cones of the retina. The sequential effector molecules of the G_t-proteins are cGMP-specific phosphodiesterases.

Perception of bitter taste can take place via α-subunits of the G_i class; the α-subunit of these G proteins is known as *gustducin* and is highly homologous with transducin. The corresponding receptors are just beginning to be characterized (review: Margolskee, 2002). A phosphodiesterase with specificity for cyclic nucleotides and a cyclic nucleotide-gated ion channel have been identified as downstream components of the signaling cascade. Signal transmission evidently takes place here in a similar way to the vision process.

Apart from a few exceptions (G_z), the members of the G_i family are characterized by inhibition by pertussis toxin (see Section 5.5.2).

G_q Subfamily

The members of the G_q subfamily are not modifiable by pertussis toxin or cholera toxin. The signal protein next in the reaction sequence is generally the β-type of phospholipase C.

G_{12} Subfamily

The G_{12} subfamily has been implicated in such cellular processes as reorganization of the cytoskeleton, activation of the c-Jun N-terminal protein kinase (see Section 10.2.2), and stimulation of Na^+/H^+ exchange. Activation by thromboxane and thrombin receptors has been described for members of the G_{12} subfamily. Examples of effector molecules include nucleotide exchange factors for Rho proteins and the GTPase-activating protein Ras-GAP (see Section 9.5).

The many α-, β- and γ-subunits now known present a large number of possibilities for subunit combination in heterotrimers. With the subunits identified at present, close to 1000 different heterotrimers could theoretically be formed. Although only a fraction of these exist in the cell, this nevertheless emphasizes the enormous complexity and diversity of regulation by G proteins.

5.5.2
Toxins as Tools in the Characterization of Heterotrimeric G Proteins

Two bacterial toxins, namely pertussis toxin and cholera toxin, were of great importance in determining the function of G proteins. Both toxins catalyze ADP ribosylation of proteins. During ADP ribosylation, an ADP-ribose residue is transferred from NAD^+ to an amino acid residue of a substrate protein (Fig. 5.17).

Cholera toxin catalyzes the ADP-ribosylation of an arginine residue (Arg174 in $G_{\alpha t}$, Arg201 in $G_{\alpha s}$) in various α-subunits. The Arg174 residue of $G_{\alpha t}$ contacts the phosphate group of the bound GTP and is thus directly involved in GTP binding and possibly also in GTP hydrolysis. Modification of Arg174 by ADP-ribosylation interferes with this

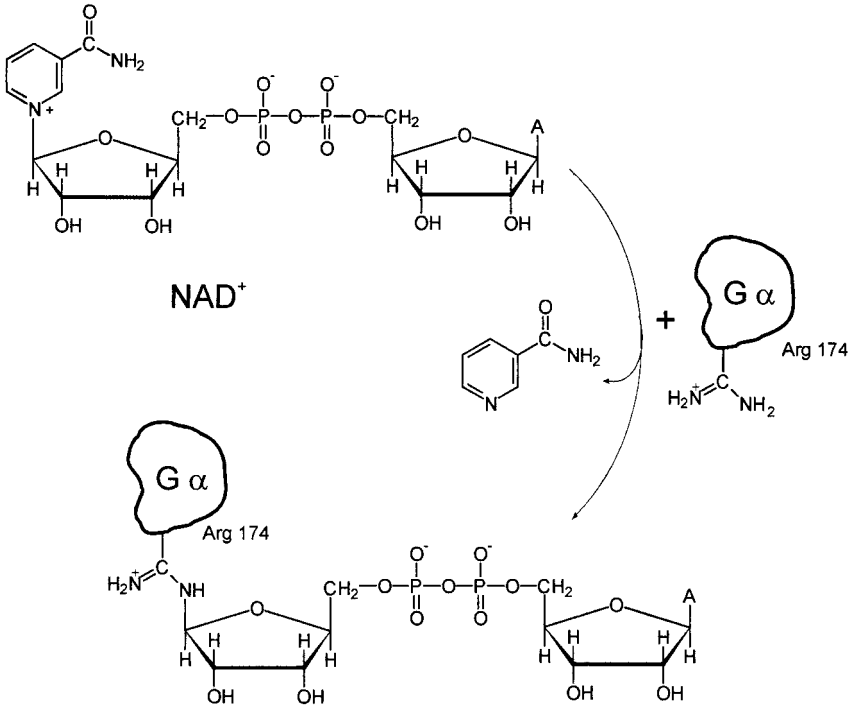


Fig. 5.17 ADP-ribosylation of the G_α-subunit of transducin by cholera toxin. Cholera toxin catalyzes the ADP-ribosylation of the α-subunit of the G-protein transducin. During the reaction, the ADP-ribose residue of NAD^+ is transferred to Arg174 of $G_{\alpha,t}$, which inactivates the GTPase activity of $G_{\alpha,t}$.

function and inactivates the GTPase activity of the G protein G protein. Consequently, the intrinsic deactivation mechanism of the G_s-protein is suspended. The G protein G protein is constitutively activated; the downstream effector molecules are – without any hormonal stimulation – permanently activated.

Constitutive activation of G_s-proteins by cholera toxin is the cause of the devastating effect of the cholera bacterium, *Vibrio cholerae*, on the water content of the intestine. Because of the lack of deactivation of the G_s-protein, adenylyl cyclase next in the reaction sequence is constantly activated, so that the level of cAMP in the cells of the intestinal epithelium is greatly increased. This, in turn, leads to increased active transport of ions, and an excessive efflux of water and Na^+ takes place in the intestine.

Pertussis toxin, formed by *Bortedella pertussis*, the causative organism of whooping cough, carries out an ADP-ribosylation at a cysteine residue close to the C-terminus of α-subunits. The modification prevents activation of the G protein G protein by the receptor, whereby the signal transmission is blocked.

5.5.3 The Functional Cycle of Heterotrimeric G Proteins

Signal transmission via G proteins takes place in close association with the inner side of the cell membrane. Both the α-subunit and the $\beta\gamma$-complex are associated with the membrane via membrane anchors (see Section 5.5.6).

Like all regulatory GTPases, the heterotrimeric G proteins run through a cyclical transition between an inactive, GDP-bound form and an active, GTP-bound form. Thereby, the activated G protein-coupled receptor functions as a nucleotide exchange factor, GEF. Figure 5.18 sketches the different functional states and the role of the individual subunits.

Inactive Ground State

In the inactive ground state, the G proteins exist as $G_\alpha \cdot GDP \cdot (\beta\gamma)$-heterotrimers. The receptor is not occupied by the hormone; the sequential effector molecule is inactive.

Activation

Binding of extracellular signal molecules (hormones, neurotransmitters) to the G protein-coupled receptor initiates activation of the G protein. The activated receptor associates with the α-subunit of the heterotrimeric complex $G_\alpha \cdot GDP \cdot (\beta\gamma)$ and induces a conformational change that leads to dissociation of GDP. It is assumed that cytoplasmic structural elements of the receptor interact with the C-terminal tail of G_α and induce an extensive conformational change that propagates to the nucleotide-binding site and leads to dissociation of GDP. The heterotrimer is now in an "empty" state, in which it possesses high affinity for the activated receptor. The free nucleotide-binding site is immediately occupied by GTP, since GTP exists in large excess compared to GDP in the cell and since the G_α-subunit binds GTP more strongly than GDP.

By catalyzing the expulsion of the bound GDP from the $G_\alpha \cdot GDP \cdot (\beta\gamma)$ complex, the activated receptor functions as a nucleotide exchange factor, GEF. It is the agonist-

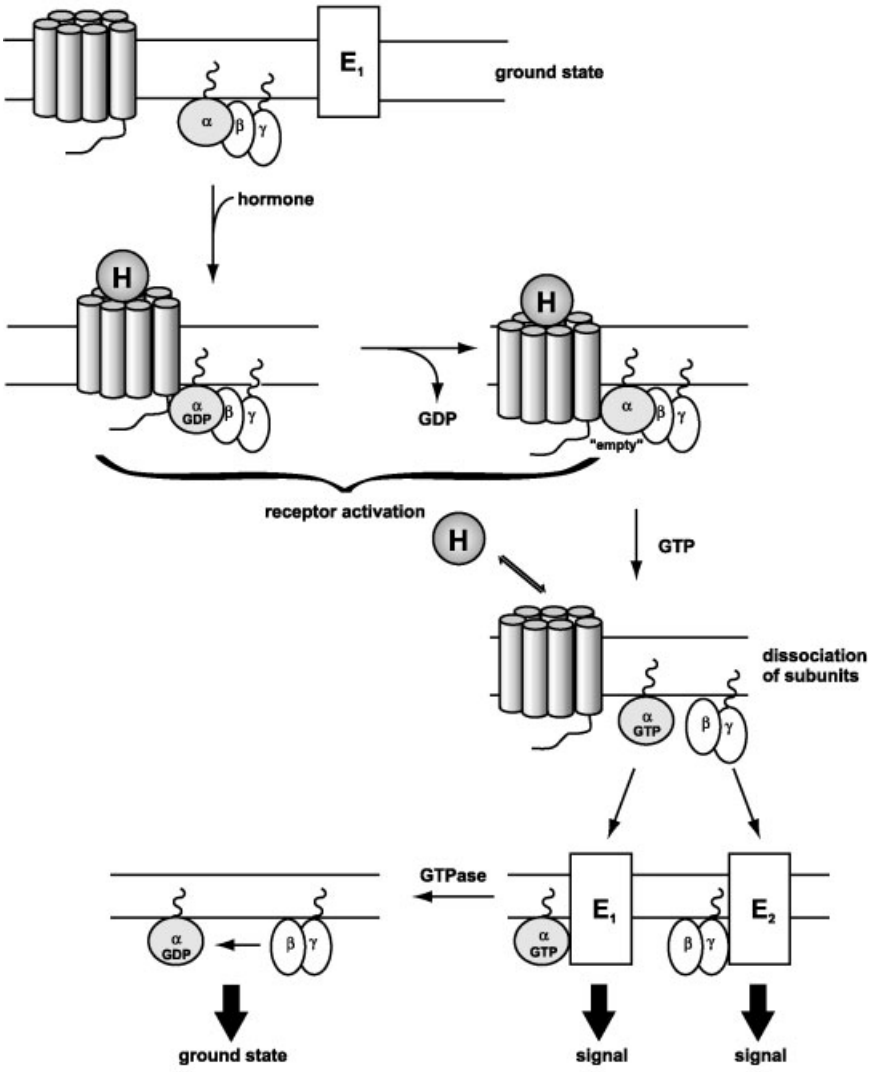


Fig. 5.18 Functional cycle of the heterotrimeric G-proteins. a) The G-proteins exist in the ground state as a heterotrimeric complex ($G_\alpha GDP) \cdot (\beta\gamma$). b) The activated receptor binds to the inactive heterotrimeric complex of the G-protein and leads to dissociation of the bound GDP and the $\beta\gamma$-complex. c) Binding of GTP to the "empty" G_α-subunit transforms the latter into the active $G_\alpha GTP$ state. $G_\alpha GTP$ interacts with an effector molecule in the sequence E1 and activates the latter for further signal transmission. The released $\beta\gamma$-complex may also take part in signal conduction by binding to a corresponding effector molecule E2 and activating the latter for further signal conduction. d) Hydrolysis of the bound GTP terminates the signal transduction via the α-subunit.

bound, activated receptor that represents the incoming signal in G protein signaling. This signal is passed on by the G_α and ($\beta\gamma$) subunits to downstream effectors and targets of the signaling cascade.

GTP Binding

GTP binding has two consequences: firstly, the $\beta\gamma$-complex dissociates and secondly, the binding to the activated receptor is cancelled. The free α-subunit with bound GTP represents the activated $G_\alpha \cdot$ GTP form of the G protein and transmits the signal further. The receptor released from the complex can activate other G proteins, enabling amplification of the signal. For a detailed model of G_α activation see Iiri et al., (1998).

Transmission of the Signal

The interaction of $G_\alpha \cdot$ GTP with the corresponding effector molecule initiates the next step in the signal transmission chain. The $\beta\gamma$-complex released during activation can also perform a signal-mediating function (see Section 5.5.5).

Termination of the Signal and GTPase-activating Proteins

Hydrolysis of GTP by the intrinsic GTPase activity of the α-subunit ends signal transmission at the level of the G proteins. The rate of GTP hydrolysis functions as an inner clock for signal transmission; it determines the lifetime of the activated state and the extent of the reactions next in sequence. The intrinsic GTPase activity of $G_\alpha \cdot$ GTP is rather low and is accelerated by the action of GTPase-activating proteins (GAPs), which thereby mostly act as negative regulators of G protein signaling.

Two classes of proteins participate in GTPase activation: specific GAPs, named regulators of G protein signaling, RGS (see Section 5.5.7), and effector proteins of the G protein signaling cascade. The RGS proteins can increase the GTPase activity by more than one order of magnitude, thereby attenuating signal transmission by G proteins. In addition, downstream effector proteins can act as GAPs. Examples include phospholipase cβ, which stimulates the intrinsic GTPase activity of the corresponding $G_{q\text{-}11}$ by close to two orders of magnitude. A further effector molecule, a certain isoform of adenylyl cyclase, has been shown to function as a GAP for the monomeric $G_\alpha \cdot$ GTP state (Scholich et al., 1999). Another example is the cGMP phosphodiesterase, which is required for the GTPAse – stimulatory action of a specific RGS protein in the vision process.

5.5.4
Structural and Mechanistic Aspects of the Switch Function of G Proteins

The reaction cycle of the heterotrimeric G proteins involves the formation and breaking of numerous protein-protein contacts. In a dynamic way, protein-protein interactions are formed and resolved during the cycle, defining distinct states of the G protein and leading to new functions and reactions. A wealth of structural information is now available for most of the distinct functional states of the G proteins. Structures are available for

- a G protein-coupled receptor in the inactive ground state: see Fig. 5.3
- G_α subunits in the inactive GDP state, the active GTP state and a presumed transition state-complex of GTP hydrolysis
- the $\beta\gamma$-complex
- G_α in complex with $\beta\gamma$
- G_α in complex with the effector cyclic GMP phosphodiesterase.

The structural data, together with a large number of biochemical, genetic and biophysical data, provide a detailed mechanistic and structural picture of how the various partners talk to each other and how the switch function of the heterotrimeric G proteins is to be explained. Only some of these points will be discussed in the following.

A structural model of the trimeric G protein and the receptor is presented in Fig. 5.19. In this model, the known structures of the ground state of rhodopsin and the structures of the transducin $Gt_\alpha \cdot GDP \cdot (\beta\gamma)$ complex have been modeled, taking into account the location of the lipid anchors and the known interaction sites between the receptor and the G protein (Hamm, 2001).

The G proteins are enzymes that can exist in different conformations, and that can undergo regulatory interactions with different partners. For a detailed understanding of the different functions of G proteins we need to answer the following questions:

- How does the activated receptor catalyze the dissociation of GDP from the $G_\alpha \cdot GDP \cdot (\beta\gamma)$ complex, and how is this reaction coupled to the dissociation of the $\beta\gamma$-complex?
- Which mechanism is the basis of GTPase activity and which residues of the α-subunit are involved in catalysis?
- What is the structural difference between the active and inactive state of the G_α-subunit and what is the structural basis for the preferential interaction of downstream effectors with the activated Gα-GTP-state?

Coupling of the Activated Receptor to the G Protein

How an activated receptor activates the downstream G protein has been mainly inferred from genetical, biochemical and biophysical experiments on the active and inactive conformation of rhodopsin and its interactions with the G protein, transducin. The data suggest that receptor activation by ligand binding causes changes in the relative orientation of the transmembrane helices 3 and 6 of rhodopsin. These changes are then thought to affect the conformation of the intracellular loops of the receptor, thereby uncovering previously masked G protein binding sites on the second, third and fourth intracellular loops. As illustrated in the model shown in Fig. 5.19, the heterotrimeric G protein interacts with the receptor via elements of the Gα-subunits and via parts of the γ-subunit. The orientation of the $G_\alpha \cdot GDP \cdot (\beta\gamma)$ complex with respect to the membrane is thought to be dictated by the lipid anchors that are found on the N-terminus of the α- and the C-terminus of the γ-subunits (see Section 5.5.6). Since the cytoplasmic surface of the ground state of rhodopsin is smaller than the interacting surface of the G protein, it is assumed that major conformational changes in both proteins are required to produce the active complex.

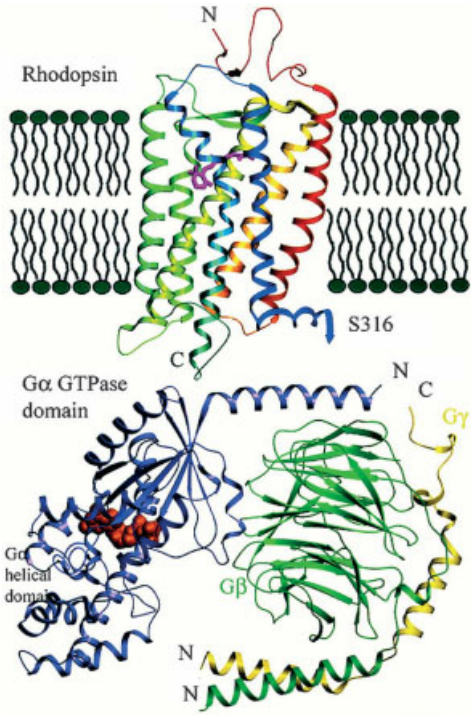


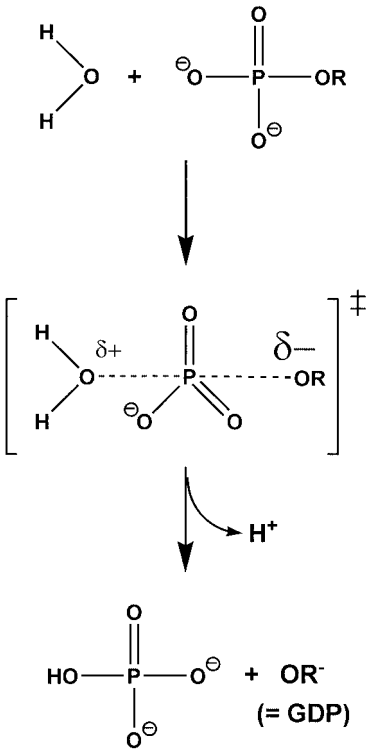
Fig. 5.19 Model of the assembly of rhodopsin and transducin at the cell membrane (from Hamm, 2001). Models are based on the crystal structure and are to scale. The C-terminal residues after S316 are not shown. The orientation of transducin ($G_{t,\alpha}$, $G_{t,\beta\gamma}$) with respect to rhodospin and the membrane is based on the charge and hydrophobicity of the surface, the known rhodosin binding sites on transducin and the sites of lipidation of $G_{t,\alpha}$, and $G_{t,\beta\gamma}$.

Mechanism of GTP Hydrolysis

The G_α-subunit possesses a slow GTPase activity, which is synonymous with a long lifetime of the activated GTP state. For transducin $G_{\alpha t}$, the k_{cat} value for GTP hydrolysis is in the region of 0.05 s^{-1}. This value is, however, close to two orders of magnitude higher than that of the Ras protein (see Chapter 9). The k_{cat} for GTP hydrolysis may be increased by up to two orders of magnitude by GTPase-activating proteins (GAP, see below) for some G_α-proteins. A much larger increase, namely 10^5-fold, is observed for the GAP belonging to the Ras protein (see Chapter 9).

It is generally assumed that hydrolysis of the γ-phosphate bond proceeds via an S_N2 mechanism, as shown in Fig. 5.20a. The hydrolysis proceeds by an "in-line" attack of a water molecule on the γ-phosphate, in which the GDP residue is displaced from the γ-phosphate. Experimental data and theoretical considerations suggest a transition state where negative charges develop mainly on the bridging $\beta\gamma$-oxygen and on the β-phosphate. In the postulated transition state, the γ-phosphorus atom is penta-coordinated, whereby the ligands are configured in the form of a trigonal bipyramid. Catalysis is thought to proceed mainly via a stabilization of the leaving group, the β-phosphate of GDP.

Information on the residues of G_α- involved in stabilization of the transition state has been obtained first from structures of G_α-proteins in complex with GDP · AlF_4^-, and a comparison of this structure with the structure of G_α · GTPγS, the activated α-subunit. AlF_4^-, has been discovered as an activator of GDP-bound α-subunits and, because of to



b)

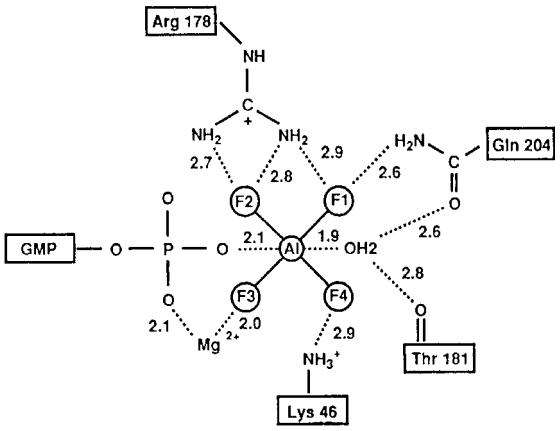


Fig. 5.20 A) "In-line" attack and dissociative mechanism of GTP hydrolysis. Hydrolysis of GTP takes place via an "in-line" attack of a water molecule at the γ-phosphate. The reaction passes through a transition state in which the γ-phosphate adopts a metaphosphate-like, planar configuration. The γ-phosphate, the water molecule and the leaving group GDP are oriented in form of a trigonal bipyramide, with an asymmetric charge distribution. A surplus of negative charge is located on the leaving group GDP and stabilizing this charge by positive residues of the protein will enhance the rate of GTP hydrolysis. This dissociative mechanism of GTP hydrolysis is now widely assumed to be used by regulatory GTPases

B) Binding of GDP · AlF_4^- at the active site of $G_{i\alpha}$. The representation is based on the structure of the $G_{i\alpha}$ · GDP AlF_4^- complex. The coordination sphere and the distances from AlF_4^- (in Å) to Arg178, Gln204, the Mg^{2+} ion and other residues of the GTPase center are shown. The catalytically important residues Gln204 and Thr181 fix a water molecule that is located "in line" to an oxygen atom of the γ-phosphate of GDP.

this characteristic – in addition to the bacterial toxins mentioned above – is often used for detection of G proteins and for their structural characterization. In the presence of AlF_4^-, permanent activation of the G protein is observed: $G_{\alpha^*}GDP$ is fixed by binding of AlF_4^- in a conformation that permits activation of the effector molecule. $GDPAlF_4^-$ functions as a *transition state analog*, in which the AlF_4^- adopts the position of the γ-phosphate in the supposed transition state of GTP hydrolysis (Fig. 5.20b). The structures of GDP AlF_4^- bound α-subunits have helped to identify the residues of G_α-proteins involved in stabilization of the transition state. It is expected from residues that participate in transition state stabilization that they assume a conformation in the transition state that is different from that of the ground state. The structure shows that the AlF_4^- is bound at the position of the γ-phosphate of $G_{i\alpha} \cdot GTP\gamma S$, such that there are noticeable differences for two residues of the protein, namely Gln204 and Arg178, between the GTP structure and the $GDP \cdot AlF_4^-$ structure..

An essential role of the Gln and Arg residues for catalysis is suggested from these structures and has also been supported by many mutational studies. A detailed real-time analysis of the GAP-catalyzed GTPAse reaction of Ras-proteins (Allin et al., 2001) supports a slightly different mechanism of the transition state stabilization as suggested from the $G_\alpha \cdot GDP \cdot AlF_4^-$ structure. According to these data, GTP hydrolysis occurrs by a dissociative mechanism (see Fig. 5.20a) whereby stabilization of the transition state is achieved mainly by interaction the Arg residue with the negative charge on the leaving group, the β-phosphate of GDP (see Section 9.2.2).

Acceleration of GTP Hydrolysis by GTPase Activating Proteins

In all, the α-subunits of the G proteins possess a slow GTPase activity. A reduction of the lifetime of the activated $G_\alpha \cdot GTP$ state, and thus weakening of the signal transmission, can be achieved by binding of specific GTPase-activating proteins such as the RGS proteins (see Section 5.5.7) to $G_\alpha \cdot GTP$. The RGS proteins stimulate the GTPase activity of different α-subunits by close to two orders of magnitude. Mechanistically, the GTPase-activating activity of the RGS proteins is explained, in particular, by stabilization of the transition state. It is assumed that the RGS proteins fix the catalytic residues of the GTPase center and bring it into a position favorable for the hydrolysis. GTPase stimulation of the Ras protein by the corresponding GAP proteins proceeds, in contrast, by another mechanism (see Section 9.2.2).

Structural Basis of the Activation of the α-Subunit

The switch function of the α-subunit of the heterotrimeric G proteins is founded on the change between an active G_α-GTP conformation and an inactive G_α-GDP conformation. In this process, interaction sites with downstream effector proteins are exposed that are not available in the inactive GDP-state. The structural difference between the two conformations was explained for the transducin, $G_{\alpha,t}$, by crystallization and structural characterization of the inactive GDP form and the active $GTP\gamma S$ form (Lambright et al., 1994). The structures of both forms of $G_{\alpha,t}$ are shown in Fig. 5.21.

$G_{\alpha,t}$ is made up of two domains, a GTPase domain and a helical domain. The *GTPase* or *G-domain* indicates that $G_{\alpha,t}$ is a member of the superfamily of regulatory GTPases. In addition, $G_{\alpha,t}$ possesses a *helical domain*, which represents a characteristic feature of

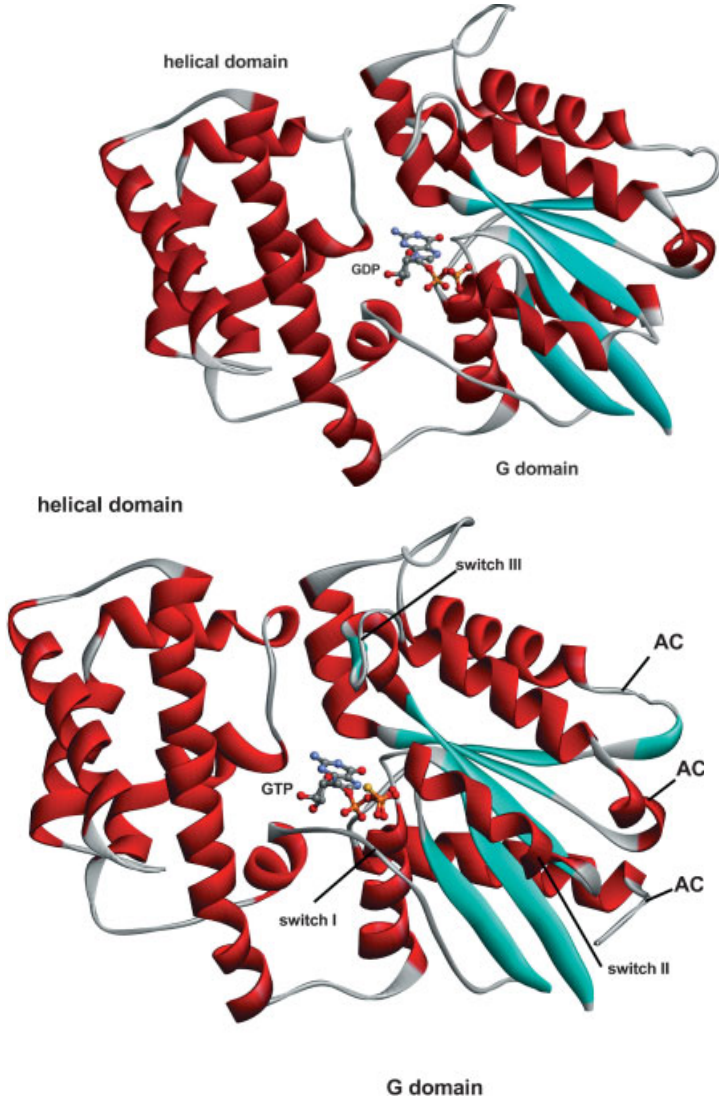


Fig. 5.21 GTP and GDP structures of Transducin. The $G_{\alpha,t}$ subunit of transducin possesses – in contrast to Ras protein and to other small regulatory GTPases – an α-helical domain that hides and closes the G-nucleotide binding pocket. The conformational changes that accompany the transition from the inactive $G_{\alpha,t}$GDP form (a) into the active $G_{\alpha,t}$GTP from (b) are restricted to three structural sections that are known as switches I, II and III. Switch I includes the link of the α-helical domain with $\beta 2$, switch II affects in particular helix $\alpha 2$, and switch III, the $\beta 3 - \alpha 3$ loop. Switch III includes a sequence that is characteristic for the α-subunits of the heterotrimeric G proteins. The conformational changes of switches II and III affect structural sections that are assumed to be binding sites for the effector molecule adenylyl cyclase (AC) and the γ-subunit of cGMP-dependent phosphodiesterase (PDEγ), based on mutation experiments and biochemical investigations. MOLSKRIPT representation according to Kraulis (1991).

the heterotrimeric G proteins. The nucleotide-binding site is in a cleft between the two domains. It is assumed that the presence of the helical domain is the reason that bound nucleotide dissociates only very slowly from transducin and that the activated receptor is therefore necessary to initiate the GDP/GTP exchange.

The γ-phosphate group of GTP must be assigned the function of a trigger of activation of $G_{\alpha,t}$. The comparison of the active and inactive conformations gives an insight into this function. In all, the active and inactive forms of $G_{\alpha,t}$ have a very similar structure. Significant conformational changes on transition between the two functional states were found for three structural elements, known as switches I, II and III, that include only 14% of the amino acids of transducin. The γ-phosphate interacts with three amino acids that move switch I upwards and thus cause a coupled movement of switches I and II (Fig. 5.21).

The GTP binding to $G_{\alpha,t}$ has several consequences:

Firstly, it is assumed that the conformational changes in switch II triggered by GTP binding lead to dissociation of the $\beta\gamma$-complex. The $\beta\gamma$-complex binds to the switch regions I and II of the α-subunit.

A further consequence of the conformational change in the α-subunit induced by the γ-phosphate is the activation of the effector molecule next in sequence. The binding site of the sequential effector molecule adenylyl cyclase includes the switch II (Tesmer et al., 1997). It is therefore assumed that the conformational change of switch II also mediates the binding and activation of the effector molecule. The binding site for the effector and for the $\beta\gamma$-complex partially overlap, so that a binding of the effector is only possible if the $\beta\gamma$-complex has dissociated (see below).

For other regulatory GTPases highly resolved structures of the active GTP form and the inactive GDP form are also available. From the comparison of the structures, the requirements and common principles of the molecular switch function of the G-domains can be defined (review: Vetter and Wittinghofer, 2001). The analogous switch elements I and II are present in other G-nucleotide binding proteins too, even though little sequence homology is visible between the analogous structural elements. The switch element III is, in contrast, characteristic for the heterotrimeric G proteins. Overall, the switch function can be described in terms of conserved conformational changes of the switch I and switch II regions triggered by a universal mechanism (Fig. 22). In the triphosphate form of the G protein, there are two hydrogen bonds from the γ-phosphate oxygens to the main chain NH groups of invariant Thr and Gly residues in switch I and switch II, respectively. The glycine is part of a conserved DXXG motif; the threonine is involved in binding Mg^{2+}. Both interactions fix the protein in a tense conformation that is released by a 'loaded spring mechanism' upon hydrolysis of the γ-phosphate allowing the transition into the relaxed GDP-specific conformation. Depending on the identity of the G-protein, the extent of the conformational change can vary from small to large changes and the size of the switch regions has to be defined for each protein.

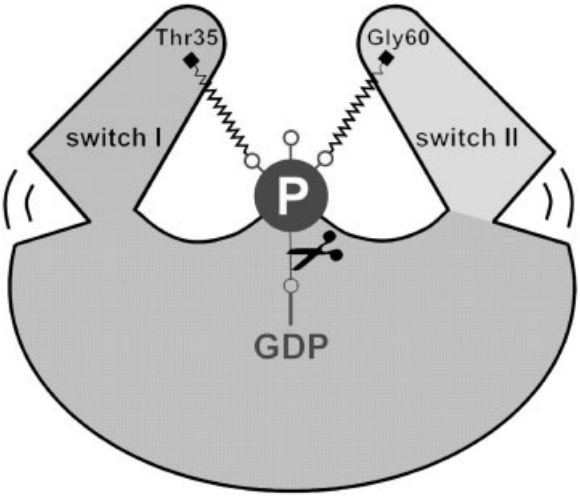


Fig. 5.22 Loaded Spring mechanism of regulatory GTPases (after Vetter and Wittinghofer, 2001) Schematic diagram of the universal switch mechanism involving switch I and switch II domains of regulatory GTPases. Invariant Thr35 and Gly60 residues (numbering for Ras, Chapter 9) are fixed in a strained conformation by H-bond interactions with the γ-phosphate of GTP. Release of the γ-phosphate after GTP hydrolysis allows the switch I and II regions to relax into the GDP-specific conformation. The extent of the conformational change is different for different proteins and may involve extra elements for some proteins.

5.5.5
Structure and Function of the βγ-Complex

Structures of the $\beta\gamma$-complex could be obtained both in the free state and in the Gα-GDP-bound state. The $\beta\gamma$-complex has an interesting configuration of seven β-sheet structures for the β-subunit. This configuration is called WD repeat or β-propeller structure because it has the form of a propellor with seven configured leaves and contains repeats of the dipeptide WD. The WD repeat motif is found in a large number of proteins and is considered a to be a stable platform that can reversibly form complexes with other proteins.

Along the GTPase cycle, the $\beta\gamma$-complex contacts at least three types of signaling proteins: the Gα-subunit, the activated receptor and the $\beta\gamma$-specific effector proteins.

Contacts to Gα are mediated only via the β-subunit, which binds in the region of the switch regions I and II and in the region of the N-terminus of the α-subunit. The binding to switches I and II of the α-subunit masks interaction sites for downstream effector proteins and prevents transmission of a signal in the ground state. The γ-subunit is located at the side of the β-subunit and does not itself interact with the α-subunit. Rather, the γ-subunit seems to be involved in the interaction with the activated receptor via its farnesylated C-terminus.

There is no great structural difference between the free and G_α-bound forms of the $\beta\gamma$-complex. Therefore, activation of the $\beta\gamma$-complex for the interaction with the corresponding effector molecule (see below) appears to be based only on its release from the inactive $G_\alpha GDP\beta\gamma$ complex. The G_α-subunit has the function of a negative regulator here, which inactivates the $\beta\gamma$-complex by masking the interaction region for signal proteins next in the sequence.

Mutational analysis has revealed the structural elements of the $\beta\gamma$-complex that interact with the effectors (Ford et al., 1998). The interaction sites with different effectors overlap partially and are only available for binding in the absence of the α-subunit. This organization of the interaction regions allows the key regulator G_α to control $G_{\beta\lambda}$ signal transmission to multiple effectors.

Originally, it was assumed that the $\beta\gamma$-complex only played a passive role in the functional cycle of the G proteins. It soon became apparent, however, that the $\beta\gamma$-complex, in addition to binding to the α-subunit, also carries out other functions and interacts specifically with corresponding effector molecules. The $\beta\gamma$-complex must be assigned its *own regulatory function*; it takes part itself in the propagation and termination of signal transmission.

The following proteins have been identified, among others, as interaction partners of $\beta\gamma$-complexes in G protein signaling:

adenylyl cyclase	phospholipase C
potassium channels	Ca^{2+}-channels
β-adrenergic receptor kinase, β-ARK	phosducin
regulators of G protein signaling, RGS	PI3-kinase type γ

The first evidence of the special function of the βγ-complex was obtained for adenylyl cyclase. Some of the adenylyl cyclases occurring in the brain are inhibited by the βγ-complex (see below); others, in contrast, are stimulated by the βγ-complex, whereby stimulation only takes place in the presence of the α-subunit.

Other effector molecules of the $\beta\gamma$-complex are specific subtypes of phospholipase C, and K^+- and Ca^{2+}-specific ion channels. In the case of Ca^{2+} channels, a direct interaction between the $\beta\gamma$-complex and cytoplasmic loops of the α-subunit of the ion channel has been demonstrated (De Waard et al., 1997). Regulation of the activity of ion channels is thus a further important role of the $\beta\gamma$-complex.

The interaction of the $\beta\gamma$-complex with G protein-coupled receptor kinases (see Section 5.3.4, β-adrenergic receptor kinase, βARK) appears to be of special regulatory importance. The function of the $\beta\gamma$-complex in this system is shown in Fig. 5.10. The $\beta\gamma$-complex binds specifically to the βARK and translocates this to the cell membrane. The translocation of βARK is necessary to switch off and modulate signal transmission via adrenaline.

5.5.6
Membrane Association of the G Proteins

Signal transmission via G proteins is inseparably linked with their membrane association. The preceding reaction partners are transmembrane proteins, and the subsquent effector molecules, such as adenylyl cyclase, are either also transmembrane proteins or are associated with the membrane (Fig. 5.23).

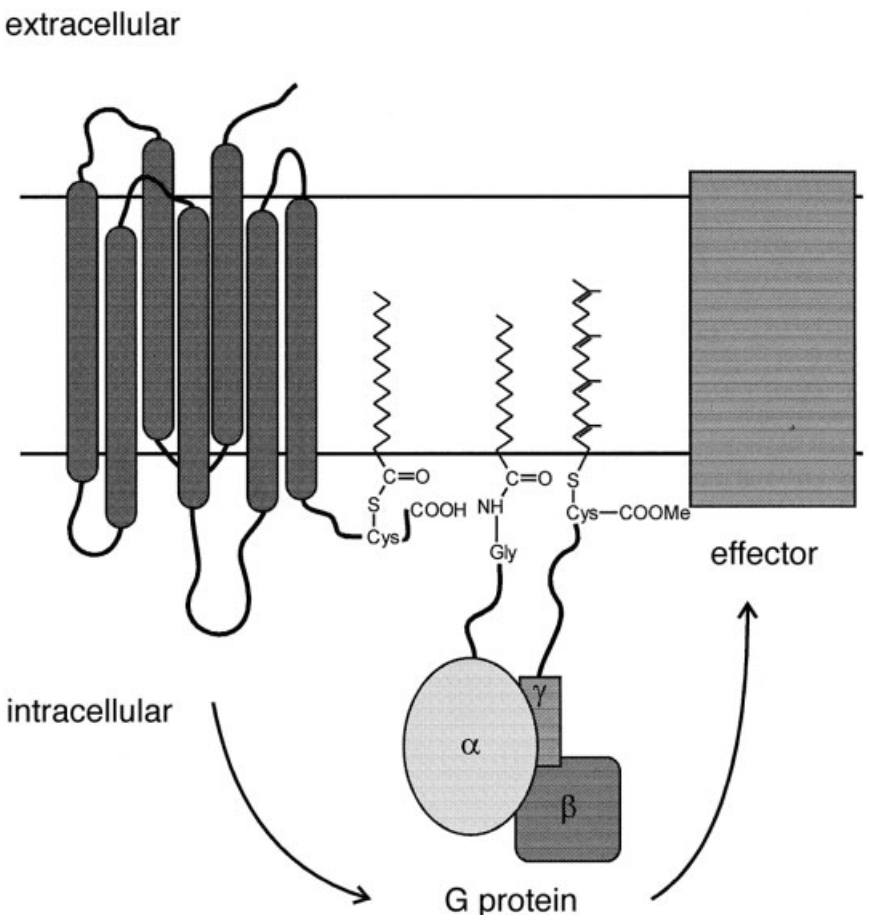


Fig. 5.23 Membrane anchor of the heterotrimeric G-proteins. The lipid anchoring in the system of G-protein-coupled receptors and the corresponding G-proteins is shown. In the figure, it is assumed that the lipid anchors are located in the membrane. A possible involvement of the lipid anchor in protein-protein interactions is not shown. The G-protein-coupled receptor carries a palmitoic acid anchor at the C-terminus. The α-subunit of the heterotrimeric G-protein is associated with the membrane via a myristoic acid anchor at the N-terminus, whilst the γ-subunit of the $\beta\gamma$-complex uses a prenyl residue as a membrane anchor.

The membrane association of the G proteins is mediated by membrane anchors that are introduced in the course of a post-translational modification at the N-terminus of the α-subunit and at the C-terminus of the γ-subunit (cf. Section 3.7).

The α-subunits of G_i and G_o subtypes possess a lipid anchor in the form of a myristoylation at the N-terminal glycine residue. The γ-subunits have a membrane anchor in the form of prenyl residues, in a similar way to Ras protein. In addition, the terminal carboxyl group is esterified with a methyl group, which further increases the hydrophobicity of the C-terminus. The length of the appended isoprenoid grouping is variable. Whilst the γ-subunit of the $G_{t,\gamma}$ protein has a farnesyl chain encompassing 15 C atoms, a modification with a C_{20} geranyl-geranyl subunit is to be found in γ-subunits of G_o-proteins in the brain.

5.5.7
Regulators of G Proteins: Phosducin and RGS Proteins

Signal transmission via G proteins and the corresponding receptors is subject to tissue- and cell-specific regulation at different levels. The regulation is mostly of a negative, suppressing character and serves two purposes in particular. Firstly, the cell must try to weaken the cytoplasmic answer under conditions of persistent activation of the receptor. Secondly, the cell needs mechanisms to rapidly terminate the signal. Typically, the rate of GTP hydrolysis of the α-subunit is very slow, about 4 min^{-1}. The cell must be able to shorten the associated long lifetime of the activated state in a regulatable way.

The most important regulatory attack points at the level of the G proteins and their receptors are (Fig. 5.24)

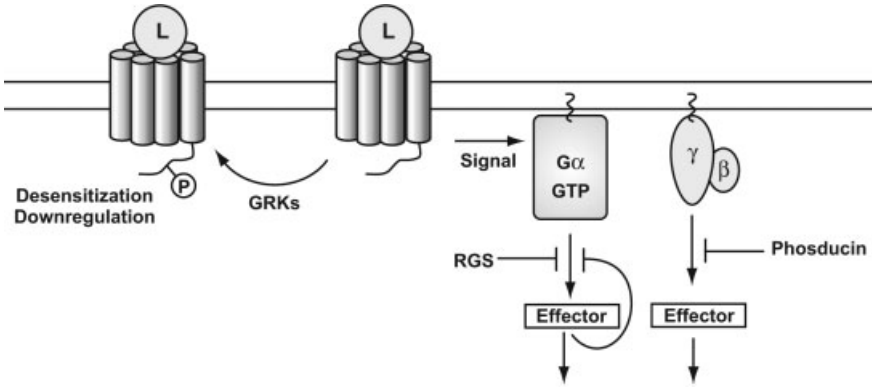


Fig. 5.24 Regulation of G protein-coupled receptors and of G proteins. Signal transduction by activated G protein-coupled receptors (GPCRs) is mainly regulated by phosphorylation via GPCR kinases (GRKs) leading to downregulation and desensitization. Signaling by $G_\alpha GTP$ can be negatively controlled by regulators of G protein signaling (RGS) and by the effector proteins themselves. A negative control of $\beta\gamma$-complex signaling is mediated by phosducin.

- Desensitization: phosphorylation of the receptor on the cytoplasmic side (see Section 5.3.4) as a reaction to persistent stimulation. This is a long-term adaptation.
- Downregulation of the number of receptor molecules: regulation at the levels of expression, stability and internalization (see Chapters 1 and 2) of the receptor.
- Inactivation of the $\beta\gamma$-complex: binding of phosducin to the $\beta\gamma$-complex.
- Reduction of the lifetime of the G_{α^*}GTP complex: activation of the GTPase of G_{α^*}GTP by RGS proteins.

At the level of the G proteins, negative regulation by phosducin or RGS proteins stands out in particular.

Regulation by Phosducin

Phosducin is an abundant protein in photoreceptor cells of the retina that is widely assumed to regulate light sensitivity through interaction with the $\beta\gamma$-subunits of the visual G protein transducin (review: Schulz, 2001). Related proteins, the *phosducin-like proteins*, are found in other tissues, e.g., in the brain and in the pineal gland. The main function of the phosducins appears to to lie in a negative regulation of signaling by G proteins. Through binding to the $\beta\gamma$-complex, the phosducins are assumed to sequester this complex and to prevent formation of the functional transducin G$\alpha\beta\gamma$ complex. In photoreceptor cells, phosducin is assigned an important role as a negative regulator, in the sense of long-term adaptation in signal transmission during the process of vision. Binding of phosducin to the β-subunit of the $\beta\gamma$-complex leads to the translocation of $\beta\gamma$ from the membrane into the cytosol. In this way, the number of $\beta\gamma$-complexes available for the G protein cycle is reduced, and the light-induced signal transmission is weakened.

Interestingly, the phosducin function is subject to regulation by phosphorylation through various protein kinases. Both protein kinase A and the Ca^{2+}/calmodulin-dependent kinase II, CaM kinase II, have been found to specifically phosphorylate Ser residues of phosducins. Phosphorylation releases the phosducin from the complex with $\beta\gamma$. The affinity for the $\beta\gamma$-complex is greatly reduced, allowing formation of the transducin heterotrimer and amplification of the visual signal. The Ser-phosphorylated phosducins are found in the cytosol in complex with proteins of the 14-3-3 family (see Section 8.2.5), which are known to recognize and bind phospho-serine residues.

Regulation by RGS Proteins

The RGS proteins have the function of GTPase-activating proteins (review: Berman and Gilman 1998; Ross and Wilkie, 2001). They bind specifically to α-subunits and activate their GTPase rate by close to two orders of magnitude. A specific domain, termed RGS domain, which is conserved among the members of the RGS family, mediates the contacts to the α-subunit. The stimulatory action of the RGS proteins on the GTPase activity of the α-subunits can be described in terms of a stabilization of the transition state of GTP hydrolysis. In the α-subunit-RGS complex, the RGS domain interacts with the switch regions of the α-subunit via multiple contact points. Importantly, a stabilization of switch II is observed, and it is assumed that this effect fixes switch II in a catalytically active conformation.

At present, at least 20 different genes are known in mammals for RGS proteins. Specific assignment to particular α-subunits is to be assumed, whereby most of the known RGS proteins act as GAPs towards members of the G_i subfamily. No RGS proteins have been found to date that are specifically directed against G_s-subunits. The GAP function appears to be exerted in this case by adenylyl cyclase isoforms. A large number of regulatory influences on RGS activity have been reported (see Wilkie and Ross, 2001). Among these are targeted ubiquitination and proteolysis, phosphorylation of RGS and their target G_α-subunits, modulation by palmitoylation and interaction with the $\beta\gamma$ complex. For the most part, it is still unclear how the different regulatory influences are integrated into G protein signaling pathways.

5.6 Effector Molecules of G Proteins

Activated G proteins pass the signal on to subsequent effector molecules that have enzyme activity or function as ion channels (see Fig. 5.16). Important effector molecules are adenylyl cyclase, phospolipases, and cGMP-specific phosphodiesterases. The activation of these enzymes leads to concentration changes of diffusible signal molecules such as cAMP, cGMP, diacylglycerol or inositol triphosphate, and Ca^{2+}, which trigger further specific reactions (see Chapter 6 and Fig. 6.1). G protein-mediated opening of ion channels may lead to changes in membrane potential and to changes in the ion environment, where the changes in Ca^{2+} concentration are of particular importance. Another effector molecule which provides a link to the signaling pathway of small GTPases of the Rho-family (see Section 9.1), is a G nucleotide exchange factor (GEF) specific for the Rho GTPases. The Rho-specfic GEF p115 is activated by G_{12}-subunits and thereby allows a crosstalk between the G protein and Rho signaling pathways.

5.6.1 Adenylyl Cyclase and cAMP as "Second Messenger"

The adenylyl cyclases catalyze the formation of 3'-5'-cyclic AMP (cAMP) from ATP (Fig. 5.25). cAMP is a widespread signal molecule that primarily functions via activation of protein kinases (see Section 6.1 and Chapter 7). Synthesis of cAMP by adenylyl cyclase is opposed by degradation and inactivation by phosphodiesterases.

Structure of Adenylyl Cyclase

Despite the central importance of adenylyl cyclase for hormonal signal transduction, its structural and functional characterization is incomplete. In mammals, at least 9 different membrane-bound types of adenylyl cyclase are described; these are known as adenylyl cyclases of types I – IX and show a high degree of sequence homology (ca. 50%) (review: Tesmer and Sprang, 1998; Patel et al., 2001). In addition, one cytosolic adenylyl cyclase has been described to date.

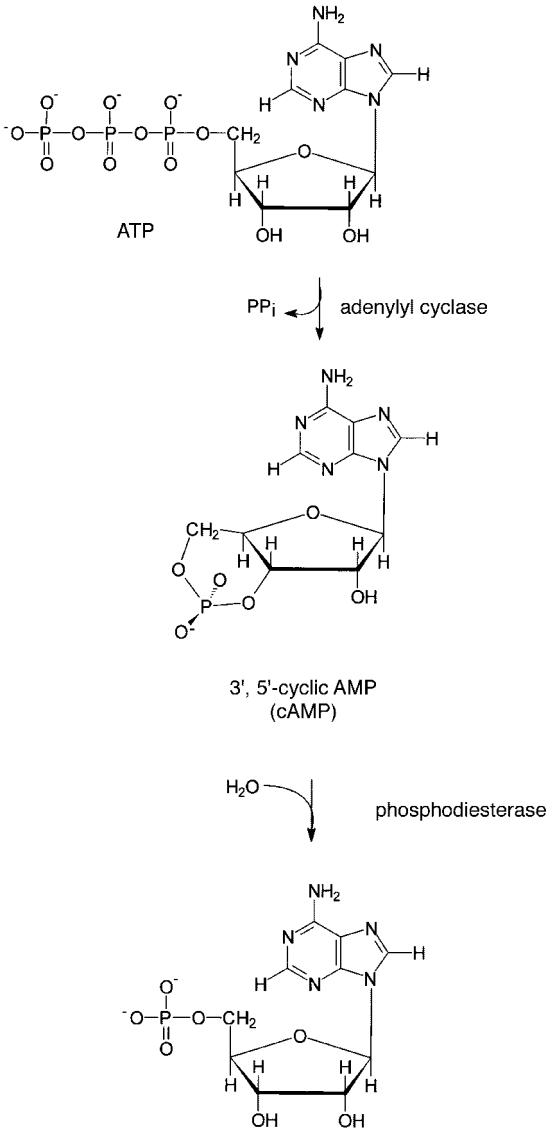


Fig. 5.25 Formation and degradation of cAMP.

The adenylyl cyclases are large transmembrane proteins with a complex transmembrane topology. The assumed topology (Fig. 5.26) shows a short cytoplasmic N-terminal section followed by a transmembrane domain M1 with six transmembrane sections, and a large cytoplasmic domain C1. The structural motif is repeated so that a second transmembrane domain M2 and a second cytoplasmic domain C2 can be differentiated.

Information on the structure-function relationship of adenylyl cyclase is available, in particular for the cytoplasmic domain. According to this, the important functions of

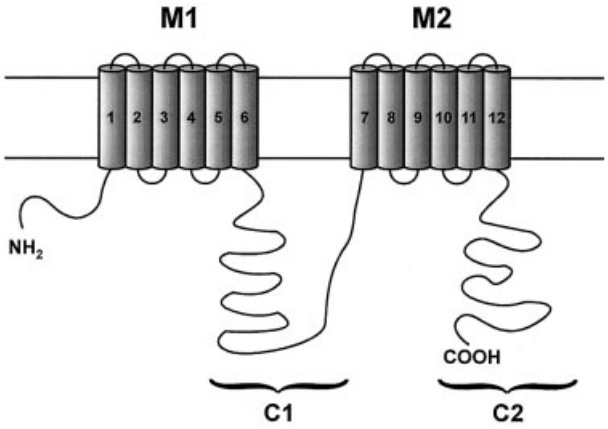


Fig. 5.26 Topology of adenylyl cyclase. The adenylyl cyclase of mammals is a transmembrane protein. It is composed of two homologous domains, which each have a transmembrane domain (M1 and M2) and a larger cytoplasmic portion (C1 and C2). Sequence analysis predicts 6 transmembrane helices for each of the domains (numbering from 1 - 12). The active site is formed by residues from C1 and C2.

adenylyl cyclase, namely the interaction with the G protein and the synthesis of cAMP, are localized on the cytoplasmic C1 and C2 domains. The C1 and C2 domains are homologous to a high degree between the different subtypes; the transmembrane domain, in contrast, is little conserved. Structural determination of the complex of $G_{\alpha,s}$, GTP and a C1-C2 dimer indicates that the active center is at the interface of the C1-C2 dimer (Zhang et al., 1997). The ATP binding site and a binding site for the activator forskolin are located there. Similar to DNA polymerases, adenlylyl cyclase appears to use a two-metal-ion mechanism for catalysis (Tesmer et al., 1999). The binding site for the regulator $G_{\alpha,s}$ GTP is relatively far away from the catalytic center. It is assumed that, during signal transmission by the G protein, an extensive conformational change is initiated that leads to a productive reorganization of the catalytic center at the C1-C2 interface.

Adenylyl Cyclase and the GTPase Cycle

Novel functions of adenylyl cyclases that affect the GTPase cycle of G proteins have been described in recent years, adding to the complexity of G protein regulation. A twofold influence of adenylyl cyclase V on the GTPase cycle of specific G proteins has been reported, namely a GAP function and a stimulation of receptor activity. While for most heterotrimeric G proteins, specific RGS proteins have been described, no $G_{\alpha,s}$-specific RGS could be found. The GTPase activating function for $G_{\alpha,s}$ appears to be fulfilled by adenylyl cyclase. Adenylyl cyclase isoform V has been shown to stimulate the GTPase activity of $G_{\alpha,s}$-subunits in a similar way to that of other effector proteins such as phospholipase β and the γ-subunit of cGMP phosphodiesterase.

Interestingly, adenylyl cyclase V also influences the activation of various G proteins by the activated receptor (Scholich et al., 1999). It facilitates the onset of signaling by agonist-bound receptors so that the amount of receptor required to activate the enzyme is drastically reduced. This function may serve to amplify signaling in the presence of low concentrations of ligand or receptor. The mechanistic basis of this influence remains to be elucidated. An apparent paradox arises in which the same protein inactivates $G_{\alpha,s}$ and facilitates activation of various G_s.

Regulation of Adenylyl Cyclase

A common feature of the different adenylyl cyclases is the stimulation of their enzyme activity by the GTP-bound form of the α-subunit of the G_s-protein. Furthermore, all subtypes, except subtype IX, are stimulated by the diterpene *forskolin*. In addition to the central regulation by the activated α-subunit, there are a variety of stimulatory or inhibitory influences on the different subtypes of adenylyl cyclase, in a manner characteristic for the particular subtype. The various subtypes differ in these regulatory influences, no two subtypes having identical patterns of regulation.

Stimulation of adenylyl cyclase may take place by
- $G_{\alpha,s} \cdot$ GTP
- Ca^{2+}/calmodulin
- protein kinase C
- $\beta\gamma$-subunits of G proteins.

Inhibition of adenylyl cyclase is possible by
- $G_{\alpha,i} \cdot$ GTP
- Ca^{2+}
- $\beta\gamma$-subunits of G proteins.

The different adenylyl cyclase isoforms demonstrate significant diversity in their regulation. Therefore, the various family members can be broadly divided into groups according to sequence homology and regulatory properties. Fig. 5.27 summarizes the stimulatory and inhibitory influences that take effect on the various groups of adenylyl cyclases.

The regulation of the subtypes I, III and VIII by Ca^{2+}/calmodulin stands out. All three subtypes are stimulated by Ca^{2+}, although in different concentration regions. Ca^{2+} is, as discussed in more detail in Chapter 6, a central intracellular messenger substance, and an increase in the Ca^{2+} concentration is observed on activation of different signal transduction processes.

In the brain, the Ca^{2+}/calmodulin regulation of adenylyl cyclase is of particular importance. One finds adenylyl cyclase concentrated in the vicinity of receptors for *N*-methyl-D-aspartate, which represent regulatable entry points for Ca^{2+}. Since the entry point for Ca^{2+} and adenylyl cyclase are in the neighborhood of one another, a rapid reaction of the cyclase to changes in Ca^{2+} concentration is ensured. According to knockout studies in mice, adenylyl cyclase isoforms play an important role in the nervous system by, e.g., influencing processes of learning, memory and behavior.

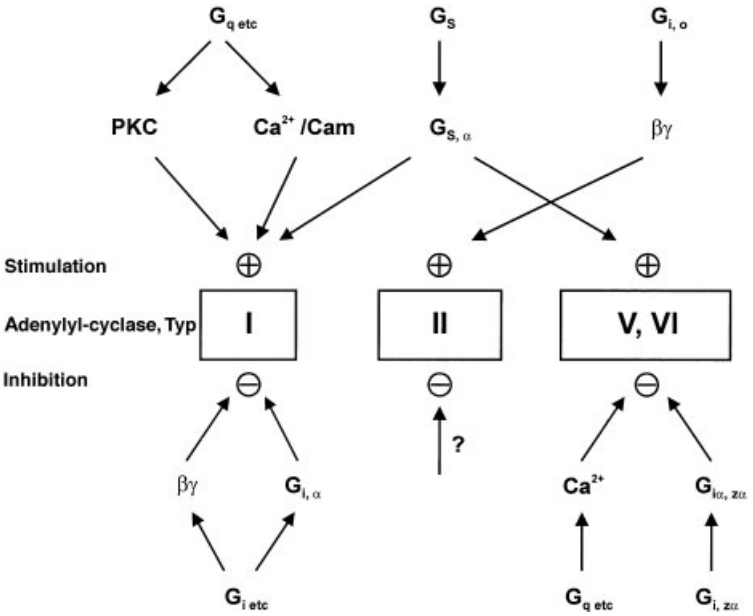


Fig. 5.27 Diversity of regulation of adenylyl cyclase. The figure summarizes schematically the regulation of adenylyl cyclases of type I, II, V and VI (after Taussig and Gilman, 1995). The individual subtypes are negatively (–) or positively (+) regulated by heterotrimeric G-proteins by various pathways. Regulation takes place both via G_α-subunits and via $\beta\gamma$-complexes. For nomenclature of the G-proteins, see 5.5.1; PKC: protein kinase C; Ca^{2+}/CaM: Ca^{2+}/calmodulin (see Chapters 6 and 7).

The cooperation of the different regulatory signals that may have an effect on adenylyl cyclase is only poorly understood. Because of the existence of different subtypes of adenylyl cyclase and their various regulation patterns, it is assumed that the individual subtypes have specific roles in different cell compartments, cells and tissues, and are each subject to defined regulatory processes.

The diverse regulation of adenylyl cyclase emphasizes the important role of this enzyme class in the signal processing in a cell. The adenylyl cyclases represent a meeting point at which different regulatory signals arrive and are weighed up against each other. In many aspects, the adenylyl cyclases are like a coincidence detector that is only activated when several signals become effective simultaneously. Ca^{2+}/calmodulin-dependent adenylyl cyclases are seen as an important element in learning processes and in memory formation. Both are processes for which a coincidence mechanism is postulated.

5.6.2 Phospholipase C

Another large class of effector molecules that are activated by G proteins is the β-subfamily of the phospholipases of type C.

Phospholipases are enzymes that cleave phospholipids. Phospholipases of type A1, A2, C and D are differentiated according to the specificity of the attack point on the phospholipid. The bonds cleaved by these phospholipases are shown in Fig. 5.28a.

Cleavage of inositol-containing phospholipids by phospholipase C is of particular regulatory importance, since this reaction generates two second messengers. Phospholipase C catalyzes the release of diacyl glycerol and inositol-1,4,5-triphosphate from phosphatidyl inositol-4,5-diphosphate, a phospholipid occurring at low concentrations in the membrane (Fig. 5.28b). Thus, phospholipase C has a key function in the formation of the intracellular messenger substances diacylglycerol, inositol-1,4,5-triphosphate and Ca^{2+} (see Chapter 6).

To date, at least 11 mammalian isoforms of phospholipase C are known. Based on sequence homology, four subfamilies β, γ, δ and ε are currently differentiated (review: Rebecchi et al., 2000). Of these, the β and γ subfamilies are key players in several signaling pathways and are best characterized. The domain structure of the phospholipases Cβ and Cγ is shown in Fig. 5.28C. Common to all phospholipases C is the occurrence of pleckstrin homology domains (PH domains). The PH domains are protein modules for which a role in mediation of protein-membrane interactions and protein-protein interactions is assumed (see Section 8.2.4). For the PH domain of phospholipase Cβ, specific binding of phosphatidyl inositol phosphates has been demonstrated. It is generally assumed that the PH domain has the function of associating the phospholipase with the membrane-localized substrate, the PtdIns(4,5)P_2, and of ensuring an effective conversion of the substrate. The catalytic activity is located on the domains X and Y. In addition, phospholipases of type C contain domains called EF-hands which mediate low-affinity binding of Ca^{2+}. Other domains found on phospholipase C are specific for the two subfamilies.

Although the phospholipases of type Cβ and Cγ catalyze the same biochemical reaction, they are activated via different signaling pathways. The Cβ subfamily participates in G protein signaling while the members of the Cγ subfamily function as effectors of receptor tyrosine kinases (see Chapter 8).

Phospholipase Cβ

Phospholipases of type Cβ function as effector enzymes in signal transmission by various G protein coupled receptors. The initiating external signals are diverse (see Fig. 5.16) and include hormones, neurohormones and sensory signals such as odorous agents and light (in nonvertebrates).

The effector function of phospholipase Cβ enzymes in G protein signaling is based on and mediated by the following functions and interactions:

– activation by the Gq subfamily of pertussis toxin-insensitive α-subunits
– activation by βγ-subunits

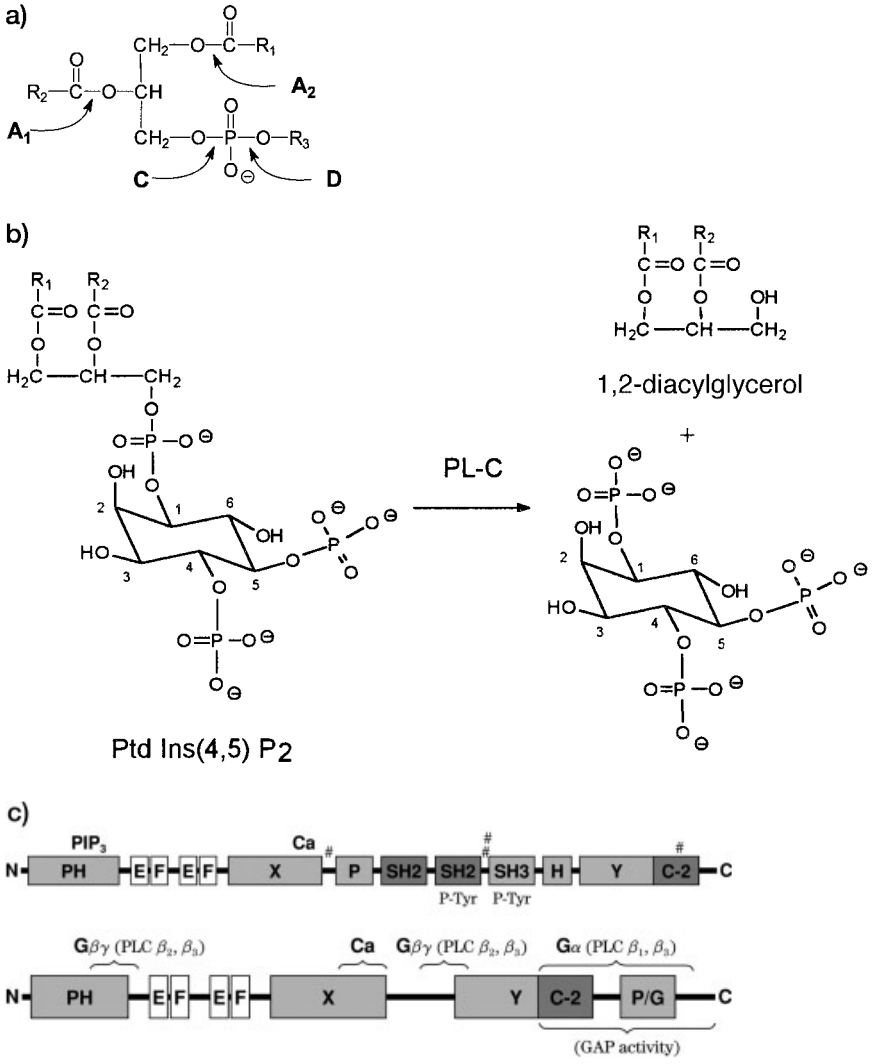


Fig. 5.28 Classification of the phospholipases and the reaction of phospholipase C. a) Cleavage specificity of phospholipases A1, A2, C and D. b) Cleavage of inositol-containing phospholipids by phospholipase C. In a reaction of particular importance for signal transduction, phospholipase C (PL-C) catalyzes the cleavage of phosphatidyl inositol-4,5-bisphosphate (PtdIns(4,5)P_2) into the messenger substances diacylglycerol and inositol 1,4,5-triphosphate (Ins(1,4,5)P_3). c) Domain structure of phospholipase Cβ and Cγ. Top: Phospholipase Cβ; bottom: phospholipase Cγ X, Y: domains required for phospholipase activity; E,F: EF-Hand; PH: pleckstrin homology domain;C2: conserved domain; P/G: C-terminal domain with phopholipid binding activity;
Proteins and ligands interacting with the domains are shown above. PIP_3: phosphatidylinositol (3,4,5) trisphosphate.

– GTPase activating function toward Gq subunits
– interaction with PDZ domain-containing proteins.

The main effector function of phospholipase Cβ enzymes is based on their stimulation by Gαq and the $\beta\gamma$-complex, whereby each of the various phospholipase Cβ isoforms has a different sensitivity to Gαq- and $\beta\gamma$-subunits. Newly discovered activities of phospholipase Cβ enzymes include a GAP activity toward Gq subunits which is thought to improve signal quality by decreasing agonist-independent noise. Furthermore, phospholipase Cβ enzymes contain regions that mediate interaction with PDZ domains of adaptor or scaffolding proteins (see Section 8.2.6). PDZ-containing proteins are involved in the clustering and structural organization of receptors and their downstream signaling proteins, and thereby regulate agonist-dependent signal transduction in, e.g., neuronal cells. It is speculated that this interaction contributes significantly to the strong association of phospholipase Cβ enzymes with membrane fractions.

The phospholipase Cβ enzymes are subject to a variety of regulatory influences. Phosphorylation by cAMP-dependent protein kinases (protein kinase A) and by protein kinase C link the β-isoforms to heterologous and homologous receptor pathways.

Phospholipase Cγ

Phospholipases of type Cγ are activated by receptor tyrosine kinases (see Chapter 8), and thus phospholipase Cγ is involved in growth factor-controlled signal transduction pathways. The receptor tyrosine kinases (see Chapter 8) phosphorylate the enzyme at specific tyrosine residues and initiate activation of the enzyme. This activation mobilizes internal calcium stores and engages multiple protein kinase pathways. Characteristic for the structure of phospholipase Cγ is the occurrence of SH2 and SH3 domains (see Chapter 8). These represent protein modules that serve to attach upstream and downstream partner proteins. The SH2 domains mediate binding to Tyr-phosphates of the activated, autophosphorylated receptor tyrosine kinase. During this process, the phospholipase Cγ enzymes are phosphorylated on Tyr-residues and are thereby activated.

The cellular function and mechanism of activation of phospholipase Cδ enzymes is only poorly characterized. Phospholipase Cδ1 contains a PH domain which by binding PtdInsP3 allows tethering to the cell membrane. A regulation by an atypical G protein, G_h, has been reported that appears to couple phospholipase Cδ1 to G protein-coupled receptors (review: Rebecchi and Pentyala, 2000).

Reference

1 Allin, C., Ahmadian, M.R., Wittinghofer, A. and Gerwert, K. (2001) Monitoring the GAP catalyzed H-Ras GTPase reaction at atomic resolution in real time. *Proc. Natl. Acad. Sci. U.S.A*, **98**, 7754–7759.

2 Angers, S., Salahpour, A. and Bouvier, M. (2002) Dimerization: an emerging concept for G protein-coupled receptor ontogeny and function. *Annu. Rev. Pharmacol. Toxicol.*, **42**, 409–435.

3 Berchthold, H., Reshetnikova, L., Reiser, C.O.A., Schirmer, N.K., Sprinzl, M. and Hilgenfeld, R. 'Crystal structure of active elongation factor Tu reveals major domain rearrangements' (1993) *Nature,* **365**, 126–132

4 Berman, D.M. and Gilman, A.G. (1998) Mammalian RGS proteins: barbarians at the gate. *J. Biol. Chem.*, **273**, 1269–1272.

5 Daaka, Y., Luttrell, L.M. and Lefkowitz, R.J. (1997) Switching of the coupling of the beta2-adrenergic receptor to different G proteins by protein kinase A. *Nature*, **390**, 88–91.

6 De Waard, M., Liu, H., Walker, D., Scott, V.E., Gurnett, C.A. and Campbell, K.P. (1997) Direct binding of G-protein beta-gamma complex to voltage-dependent calcium channels. *Nature*, **385**, 446–450.

7 Downes, G.B. and Gautam, N. (1999) The G protein subunit gene families. *Genomics*, **62**, 544–552.

8 Ford, C.E., Skiba, N.P., Bae, H., Daaka, Y., Reuveny, E., Shekter, L.R., Rosal, R., Weng, G., Yang, C.S., Iyengar, R. *et al.* (1998) Molecular basis for interactions of G protein betagamma subunits with effectors. *Science*, **280**, 1271–1274.

9 Freedman, N.J. and Lefkowitz, R.J. (1996) Desensitization of G protein-coupled receptors. *Recent Prog. Horm. Res.*, **51**, 319–351.

10 Gether, U. (2000) Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocr. Rev.*, **21**, 90–113.

11 Gschwind, A., Zwick, E., Prenzel, N., Leserer, M. and Ullrich, A. (2001) Cell communication networks: epidermal growth factor receptor transactivation as the paradigm for interreceptor signal transmission. *Oncogene*, **20**, 1594–1600.

12 Hamm, H.E. (2001) How activated receptors couple to G proteins. *Proc. Natl. Acad. Sci. U.S.A*, **98**, 4819–4821.

13 Iiri, T., Farfel, Z. and Bourne, H.R. (1998) G-protein diseases furnish a model for the turn-on switch. *Nature*, **394**, 35–38.

14 Lambright, D.G., Noel, J.P., Hamm, H.E. and Sigler, P.B. (1994) Structural determinants for activation of the alpha-subunit of a heterotrimeric G protein. *Nature*, **369**, 621–628.

15 Margolskee, R.F. (2002) Molecular mechanisms of bitter and sweet taste transduction. *J. Biol. Chem.*, **277**, 1–4.

16 McDonald, P.H. and Lefkowitz, R.J. (2001) Beta-Arrestins: new roles in regulating heptahelical receptors' functions. *Cell Signal.* , **13**, 683–689.

17 Meng, E.C. and Bourne, H.R. (2001) Receptor activation: what does the rhodopsin structure tell us? *Trends Pharmacol. Sci.*, **22**, 587–593.

18 Okada, T., Ernst, O.P., Palczewski, K. and Hofmann, K.P. (2001) Activation of rhodopsin: new insights from structural and biochemical studies. *Trends Biochem. Sci.*, **26**, 318–324.

19 Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le, T., I, Teller, D.C., Okada, T., Stenkamp, R.E. *et al.* (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, **289**, 739–745.

20 Patel, T.B., Du, Z., Pierre, S., Cartin, L. and Scholich, K. (2001) Molecular biological approaches to unravel adenylyl cyclase signaling and function. *Gene*, **269**, 13–25.

21 Pitcher, J.A., Freedman, N.J. and Lefkowitz, R.J. (1998) G protein-coupled receptor kinases. *Annu. Rev. Biochem.*, **67**, 653–692.

22 Rastinejad, F. (2001) Retinoid X receptor and its partners in the nuclear receptor family. *Curr. Opin. Struct. Biol.*, **11**, 33–38.

23 Rebecchi, M.J. and Pentyala, S.N. (2000) Structure, function, and control of phosphoinositide-specific phospholipase C. *Physiol Rev.*, **80**, 1291–1335.

24 Ross, E.M. and Wilkie, T.M. (2000) GTPase-activating proteins for heterotrimeric G proteins: regulators of G protein signaling (RGS) and RGS-like proteins. *Annu. Rev. Biochem.*, **69**, 795–827.

25 Scholich, K., Mullenix, J.B., Wittpoth, C., Poppleton, H.M., Pierre, S.C., Lindorfer, M.A., Garrison, J.C. and Patel, T.B. (1999) Facilitation of signal onset and termination by adenylyl cyclase. *Science*, **283**, 1328–1331.
26 Schulz, R. (2001) The pharmacology of phosducin. *Pharmacol. Res.*, **43**, 1–10.
27 Sunahara, R.K., Tesmer, J.J., Gilman, A.G. and Sprang, S.R. (1997) Crystal structure of the adenylyl cyclase activator Gsalpha. *Science*, **278**, 1943–1947.
28 Tesmer, J.J. and Sprang, S.R. (1998) The structure, catalytic mechanism and regulation of adenylyl cyclase. *Curr. Opin. Struct. Biol.*, **8**, 713–719.
29 Tesmer, J.J., Sunahara, R.K., Johnson, R.A., Gosselin, G., Gilman, A.G. and Sprang, S.R. (1999) Two-metal-Ion catalysis in adenylyl cyclase. *Science*, **285**, 756–760.
30 Vetter, I.R. and Wittinghofer, A. (2001) The guanine nucleotide-binding switch in three dimensions. *Science*, **294**, 1299–1304.
31 Zhang, G., Liu, Y., Ruoho, A.E. and Hurley, J.H. (1997) Structure of the adenylyl cyclase catalytic core. *Nature*, **386**, 247–253.

6
Intracellular Messenger Substances: "Second Messengers"

6.1
General Functions of Intracellular Messenger Substances

Extracellular signals are registered by membrane receptors and conducted into the cell via cascades of coupled reactions. The first steps of signal transmission often take place in close association with the membrane, before the signal is conducted into the cell interior. The cell uses mainly two mechanisms for transmission of signals at the cytosolic side of the membrane and in the cell interior. Signal transmission may be mediated by a protein-protein interaction. The proteins involved may be receptors, proteins with adaptor function alone, or enzymes. Signals may also be transmitted with the help of low-molecular-weight messenger substances. These are known as "*second messengers*". The intracellular messenger substances are formed or released by specific enzyme reactions during the process of signal transduction, and serve as effectors, with which the activity of proteins further in the sequence is regulated (Fig. 6.1).

The intracellular messengers are *diffusible signal molecules* and reach their target proteins mostly by diffusion. Close spatial proximity of the signal components, as achieved for transmembrane receptors and their effector proteins with the help of membrane anchoring or with specific protein-protein modules (see Chapters 5 and 8), is not necessarily required for this type of signal transduction.

Two types of intracellular messenger substance can be differentiated (see Fig. 6.1):

- Messenger substances with *hydrophobic character* such as diacyl glycerol or the phosphatidyl inositol derivatives are membrane localized. The hydrophobic messengers reach membrane-associated effector proteins by diffusing through the plasma membrane and there regulate their activity.
- *Hydrophilic* messengers with good aqueous solubility are localized in the cytosol and reach their protein substrates in the cytosol.

The most important "second messengers" are
– hydrophilic, cytosolic:
cAMP, cGMP
inositol phosphates
Ca^{2+}

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

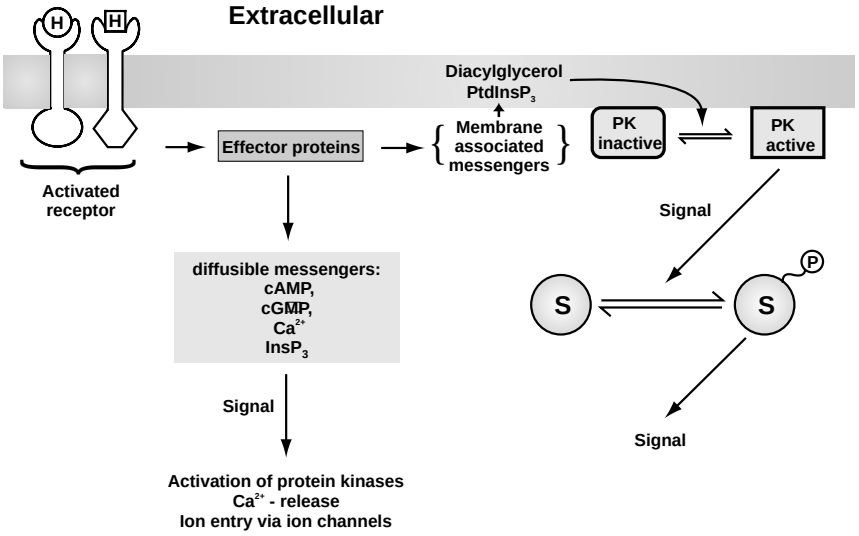


Fig. 6.1 Function and formation of intracellular messenger substances in signaling pathways. Starting from the activated receptor, effector proteins next in sequence are activated that create an intracellular signal in the form of diffusible messenger substances. The hydrophilic messenger substances diffuse to target proteins in the cytosol and activate these for signal transmission further. Hydrophobic messenger substances, in contrast, remain in the cell membrane and diffuse at the level of the cell membrane to membrane-localized target proteins. PK: protein kinase; S: substrate of the protein kinase.

– hydrophobic, membrane-associated:
  diacylglycerol
  phosphatidyl inositol phosphates.

The intracellular "second messengers" are characterized by a series of properties that make them particularly suitable as elements of signal transduction:

- Intracellular messenger substances can be formed and degraded again in specific enzyme reactions. Via enzymatic pathways, large amounts of messenger substances can be rapidly created and inactivated again.
- Messenger substances such as Ca^{2+} may be stored in special storage organelles, from which they can be rapidly released by a signal.
- Messenger substances may be produced in a location-specific manner, and they may also be removed or inactivated according to their location. It is therefore possible for the cell to create signals that are spatially and temporally limited.

6.2 cAMP

3'-5'-cyclic AMP is a central intracellular "second messenger" that influences many cellular functions, such as gluconeogenesis, glycolysis, lipogenesis, muscle contraction, membrane secretion, learning processes, ion transport, differentiation, growth control and apoptosis.

Concentration of cAMP is controlled primarily by two means, namely via new synthesis by adenylyl cyclase and degradation by phosphodiesterases (review: Houslay and Milligan, 1997). The activity of adenylyl cyclase is sensitive to signals mediated by G-protein-coupled receptors and involving Gα- and G$\beta\gamma$-proteins. In addition, cAMP production by adenylyl cyclase can be regulated via Ca^{2+}-signals (see Section 5.6.1).

Degradation of cAMP by cAMP phosphodiesterases is also an important point of attack for control of the cAMP level. More than ten different isoforms of phosphodiesterase are known, which vary in their cyclic nucleotide specificity and in their regulation. The phosphodiesterases are subject to a variety of regulatory influences, including regulation by Ca^{2+}/calmodulin and by protein phosphorylation (review: Francis et al., 2001).

cAMP functions as an activator of downstream signaling proteins which possess specific cAMP binding sites and are regulated by cAMP via allosteric mechanisms. To achieve activation of the targets of cAMP, an increase in cAMP concentration over certain threshold values is usually required. This may be reached via both paths, via stimulation of adenylyl cyclase and/or inhibition of the phosphodiesterase. Another important regulatory aspect of cAMP-mediated signal transduction is the specific localization or compartmentalization of both the synthesis and degradation of cAMP, which allows the creation of a spatially and temporally limited signal on the inner side of the cell membrane. The proteins involved in cAMP signal conduction perform their function, without exception, in association with the cell membrane.

Cyclic AMP binds to and activates the following signaling proteins:

cAMP-gated Ion Channels

An important function of cAMP is the regulation of ion passage through cAMP-gated ion channels. cAMP binds to cytoplasmic structural elements of these ion channels and regulates their open state. An example is the cAMP-regulated Ca^{2+} passage through cation channels. cAMP also performs this function during the perception of smell in mammals.

Protein Kinase A

The majority of the biological effects of cAMP are mediated by the activation of protein kinases. Protein kinases regulated by cAMP are classified as protein kinase A (see Section 7.3).

The mechanism of activation of protein kinases of type A by cAMP is schematically represented in Fig. 6.2. In the absence of cAMP, protein kinase A exists as a tetramer, composed of two regulatory (R) and two catalytic (C) subunits. In the tetrameric R_2C_2 form, protein kinase A is inactive since the catalytic center of the C subunit is blocked by the R subunit.

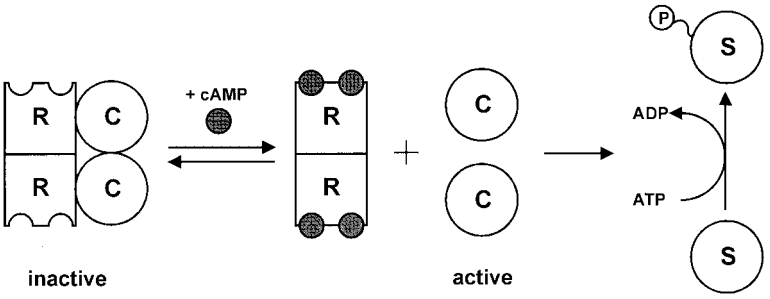


Fig. 6.2 Regulation of protein kinase A via cAMP. Protein kinase A is a tetrameric enyzme composed of two catalytic subunits (C) and two regulatory subunits (R). In the R_2C_2 form, protein kinase A is inactive. Binding of cAMP to R leads to dissociation of the tetrameric enyzme into the R_2 form with bound cAMP and free C subunits. In the free form, C is active and catalyzes the phosphorylation of substrate proteins (S) at Ser/Thr residues.

Regulation of protein kinase A by cAMP takes place by the following mechanism. An increase in cAMP concentration, triggered by activation of adenylyl cyclase and/or inhibition of phosphodiesterase, leads to cooperative binding of two molecules of cAMP to two sites on each regulatory subunit. Upon binding of four molecules of cAMP, the enzyme dissociates into an R subunit dimer with four molecules of cAMP bound and two free C subunits which are now released from inhibition by the regulatory subunits and can thus phosphorylate Ser/Thr residues on specific substrate proteins.

A large part of protein kinase A is specifically associated with the cell membrane via specific anchor proteins (see Section 7.7). In addition, a specific membrane targeting of phosphodiesterases has been reported. There is much to suggest that the formation and degradation of cAMP and activation of protein kinase A occurs at spatially restricted sites on the inner side of the cell membrane, and a localized reaction is thus initiated. This aspect of signal transduction, known as targeting, is described in more detail in Section 7.7.

The nature of the substrate proteins of protein kinase of type A is very diverse; the substrate may be, e. g., other proteins or enzymes of intermediary metabolism (see Section 7.3).

Regulation of a Guanine Nucleotide Exchange Factor (GEF)

A further second-messenger function of cAMP is the activation of GEFs (see Section 9.6), which are major regulators of the signaling function of small GTPases. Increases in the concentration of cAMP have been shown to result in the binding of cAMP to a protein termed Epac, a GEF acting on the GTPase Rap1 (see Section 9.1). This causes a conformational change leading to increased exchange activity towards Rap1 (de Rooij et al., 1998) and Rap1 activation. Activation of Rap links cAMP signaling to activation of the protein kinase B-Raf and the mitogen-activated signaling (MAPK) pathway (see Chapter 10). Signals transmitted in this way may ultimately lead to changes in transcriptional activity.

6.3 cGMP

Like cAMP, 3'-5'-cGMP is widespread as an intracellular messenger substance. Analogous to cAMP, cGMP is formed by catalysis via *guanylyl cyclase* from GTP (reviews: Wedel and Garbers, 2001; Potter and Hunter, 2001).

Although the guanylyl cyclases catalyze a similar reaction as the adenylyl cyclases, the two enzyme classes differ considerably in structure and mechanism of activation. The guanylyl cyclases can be divided into three groups according to the number of transmembrane segments. One group contains enzymes that do not contain a transmembrane segment and are referred to as soluble guanylyl cyclases. A second group contains one transmembrane segment. The members of this group are directly regulated by extracellular ligands and therefore have receptor function. A third group with more than two transmembrane segments is only poorly characterized, and its ligands or mechanism of activation is not yet known.

Guanylyl Cyclases with a Single Transmembrane Segment

The guanylyl cyclases with a single transmembrane segment function as receptors that contain an extracellular ligand-binding domain and various intracellular domains that are required for the ligand-regulated activation of the enzyme (Fig. 6.3) . As ligands for the guanylyl cyclase receptors, peptides with vasodilatory properties like the atrial natriuretic peptide have been identified. The receptor-type guanylyl cyclases are therefore also termed natriuretic peptide receptors, NPR. The receptors exist in a homodimeric transmembrane form, and its intracellular guanylyl cyclase domain is activated by peptide binding to the extracellular domains. A complicated series of reactions follow activation, which include phosphorylation of an intracellular kinase-homology domain, ATP binding and finally activation of cGMP synthesis. The mechanistic basis of guanylyl cyclase activation is unknown for these receptors.

In summary, the guanylyl cyclase receptors constitute a unique class of receptors that transmit an extracellular signal directly into the formation of an intracellular second messenger substance. In this way, important physiological processes like relaxation of blood vessels are regulated.

Soluble Guanylyl Cyclases

The soluble guanylyl cyclases exist as heterodimers and are regulated by the second messenger NO (see Section 6.10.3). A heme group that confers NO-sensitivity is bound at the N-terminus of these enzymes. NO binding to the heme group results in activation of the guanylyl cyclase activity. The second messenger function of cGMP is directed towards three targets:

- **cGMP-dependent protein kinases**
  The cGMP-dependent protein kinases are activated by cGMP binding (review: Hofmann et al., 2000) and have structural elements similar to those of protein kinase A. In contrast to protein kinase A, the regulatory and catalytic functions are localized on *one* protein chain in cGMP-dependent protein kinases. Binding of cGMP to the

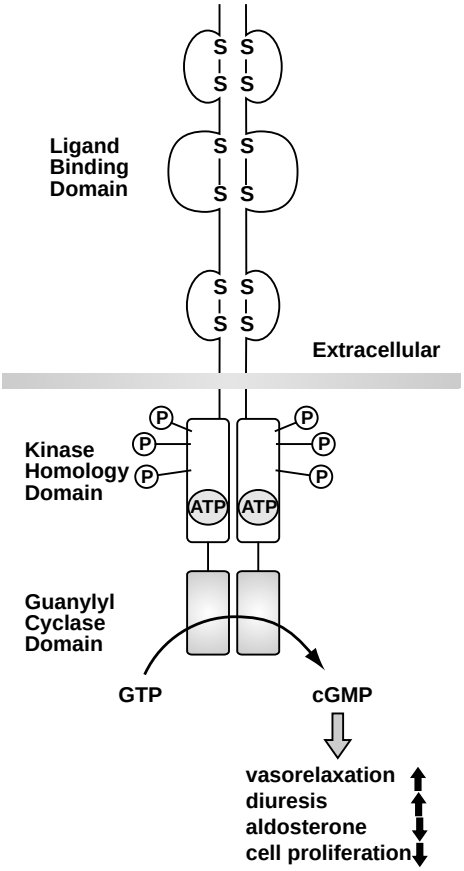


Fig. 6.3 Model of the domain structure of the natriuretic peptide receptor NPR, a receptor-type guanylyl cyclase. NPR is a dimeric transmembrane receptor which spans the membrane with two transmembrane elements. The extracytosolic domain comprises the ligand binding site and contains several disulfide bridges. The cytosolic part is composed of a kinase homology domain with multiple phosphorylation sites, an ATP binding site of unknown function and the catalytic guanylyl cyclase domain.

regulatory domain relieves autoinhibition by the N-terminus and allows phosphorylation of substrate proteins.

The functions of the cGMP-dependent protein kinases are not as well defined as those of protein kinase A. Important *in vivo* substrates in smooth muscle cells are Ca^{2+}-channels and a myosin-specific protein phosphatase. Phosphorylation of the two substrates by cGMP specific protein kinase modulates Ca^{2+}-levels and thereby controls smooth muscle tone.

- **Cation channels**
  We know of cation channels that are gated by cGMP. These channels possess cGMP-binding sites on their intracellular side. Binding of cGMP to the cation channel induces opening of the channel and the influx of cations. In the vision process, cGMP has the role of regulating Ca^{2+} influx via cGMP-gated cation channels.
- **cAMP-specific phosphodiesterases**
  Some types of cAMP-specific phosphodiesterases are regulated by cGMP.

6.4 Metabolism of Inositol Phospholipids and Inositol Phosphates

Inositol-containing phospholipids of the plasma membrane are the starting compounds for the formation of various low-molecular-weight inositol messengers in response to various intra-and extracellular signals. These messengers include the central second messengers diacylglycerol and inositol trisphosphate as well as membrane-bound phosphatidyl inositol phosphates.

The plasma membrane contains the phospholipid phosphatidyl inositol (PtdIns), in which the phosphate group is esterified with a cyclic alcohol, myo-D-inositol (Fig. 6.4). Starting from PtdIns, a series of enzymatic transformations lead to the generation of a diverse number of second messengers. PtdIns is first phosphorylated by specific kinases at the 4' and 5' positions of the inositol residue, leading to the formation of phosphatidyl inositol-4,5-bisphosphate [PtdIns(4,5)P_2].

From PtdIns(4,5)P_2, two paths lead to physiologically important messenger substances. PtdIns(4,5)P_2 may be further phosphorylated by PI3-kinase to PtdIns(3,4,5)P_3, which functions as a membrane-localized messenger (see Section 6.6). In a further reaction, PtdIns(3,4,5)P_3 may be cleaved by phospholipase C, forming the "second messengers" inositol 1,4,5-triphosphate [Ins(1,4,5)P_3] and diacylglycerol (DAG). Both compounds are messengers that can activate further specific reaction chains. Ins(1,4,5)P_3 activates the release of Ca^{2+}, whilst diacylglycerol acts primarily by stimulation of protein kinase C.

In addition to these well-characterized routes, further transformations of inositol phosphates and phosphatidyl-inositol phosphates are known which lead to the formation of nearly 30 inositol-containing compounds with potential messenger function. These reactions include phosphorylation to inositol polyphosphates as well as specific dephosphorylation by inositol phosphatases. However, for only some of these compounds the biochemical attack points are known and specific *in vivo* functions could be demonstrated (review: Irvine and Schell, 2001).

Inositol Phosphates and Regulation of Phospholipase C

Phospholipase C, which occurs in different subtypes in the cell, is a *key enzyme* of phosphatide inositol metabolism (for cleavage specificity, see Fig. 5.28). Two central signaling pathways regulate phospholipase C activity of the cell in a positive way (Fig. 6.5). Phospholipases of type Cβ (PL-Cβ) are activated by G proteins and are thus linked into signal paths starting from G protein-coupled receptors. Phospholipases of type γ (PL-Cγ), in contrast, are activated by transmembrane receptors with intrinsic or associated tyrosine kinase activity (see Chapters 8 and 10). The extracellular stimuli activated by the two major reaction pathways are very diverse in nature, which is why the phospholipase C activity of the cell is subject to multiple regulation.

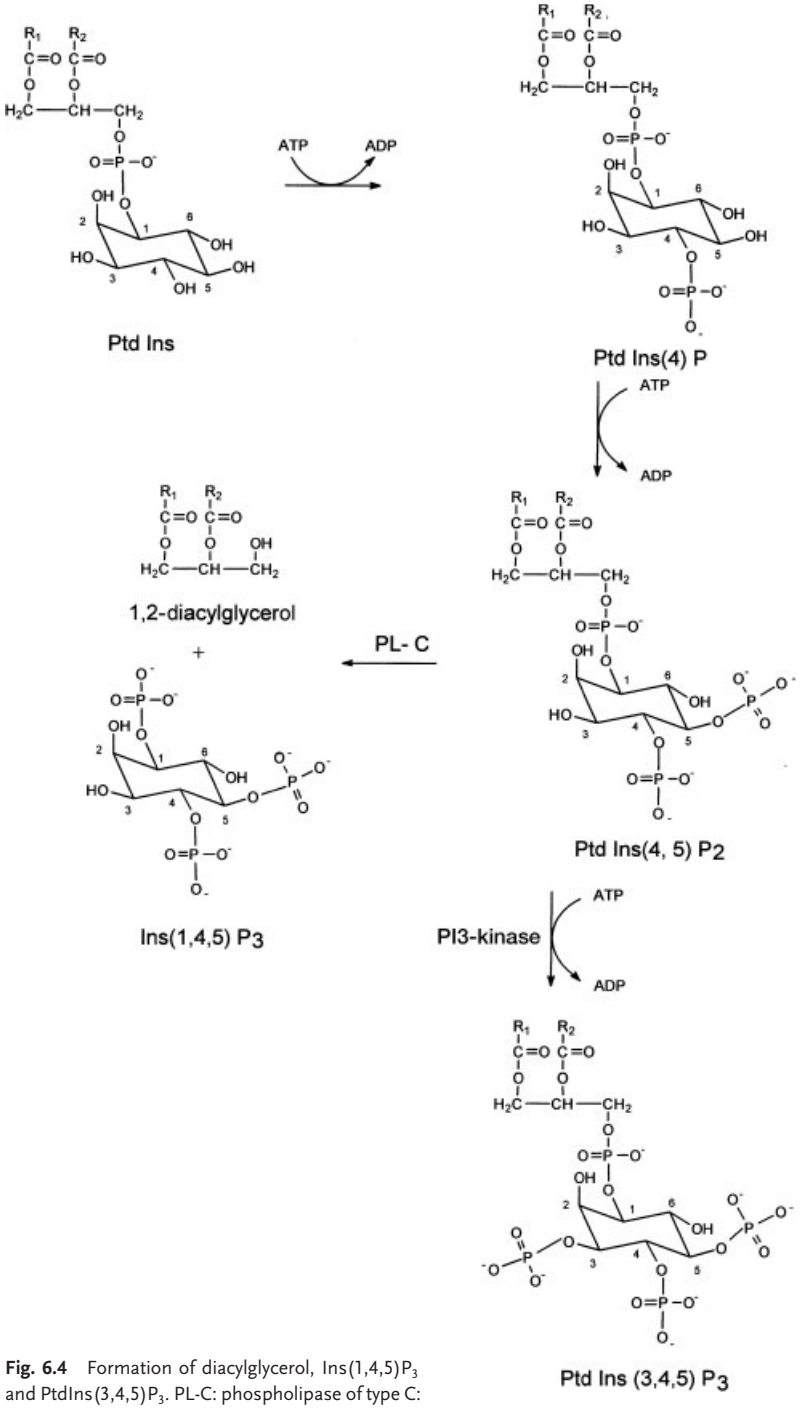


Fig. 6.4 Formation of diacylglycerol, Ins(1,4,5)P_3 and PtdIns(3,4,5)P_3. PL-C: phospholipase of type C: PI3-kinase: phosphatidyl inositol-3'-kinase.

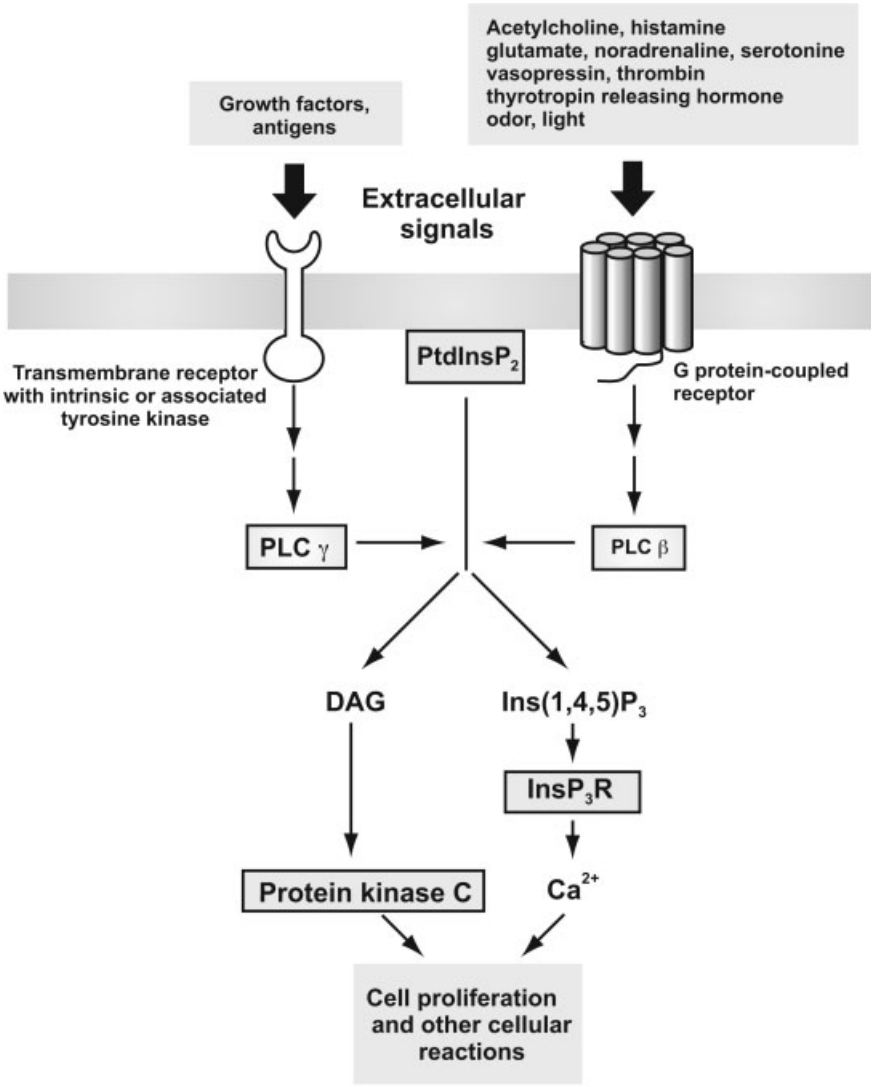


Fig. 6.5 Formation and function of diacylglycerol and Ins(1,4,5)P_3. Formation of diacylglycerol (DAG) and Ins(1,4,5)P_3 is subject to regulation by two central signaling pathways, which start from transmembrane receptors with intrinsic or associated tyrosine kinase activity (see Chapters 8 & 11) or from G-protein-coupled receptors. DAG activates protein kinase C (PKC, see Chapter 7), which has a regulatory effect on cell proliferation, via phosphorylation of substrate proteins. Ins(1,4,5)P_3 binds to corresponding receptors ($InsP_3$-R) and induces release of Ca^{2+} from internal stores. The membrane association of DAG, PtdIns(3,4)P_2 and PL-C is not shown here, for clarity.

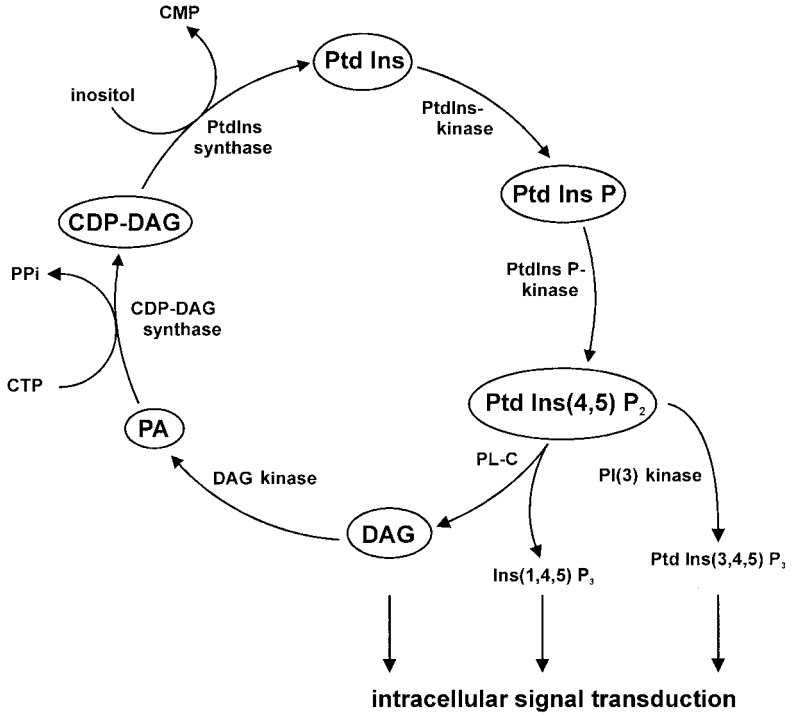


Fig. 6.6 Metabolic cycle of regeneration of PtdIns(4,5)P_2.

Metabolic Cycle of Inositol Phosphate

The inositol phosphates are linked into a metabolic cycle (Fig. 6.6) in which they can be degraded and regenerated. Via these pathways, the cell has the ability to replenish stores of inositol phosphate derivatives, according to demand. PtdIns may be regenerated from diacylglycerol via the intermediate levels of phosphatidic acid and CDP-glycerol.

Regeneration of PtdIns in the inositol cycle is of particular importance in the vision process in *Drosophila* (Wu et al., 1995). In *Drosophila*, $InsP_3$ serves as a messenger during perception of light. On incidence of light, $InsP_3$ is formed from $PtdInsP_2$. It has been shown that CDP diacylglycerol synthase, which supplies CDP diacylglycerol for the resynthesis of PtIns (see Fig. 6.6) has an essential role in light perception in *Drosophila*. If regeneration of PtdIns is not possible because of a defective CDP diacylglycerol synthase, severe disorders of the vision process occur.

6.5
Inositol 1,4,5-Triphosphate and Release of Ca^{2+}

The primary signal function of Ins(1,4,5)P_3 is the mobilization of Ca^{2+} from storage organelles. Ca^{2+} is a ubiquitous signaling molecule whose signaling function is activated by its release from intracellular stores or through Ca^{2+}-entry channels from the

extracellular side. A multitude of second messengers has been shown to induce an increase of intracellular Ca^{2+} (review: Bootman, 2002). For release from the intracellular stores, three types of messengers stand out, namely Ins(1,4,5)P_3, cyclic ADP ribose and nicotinic acid adenine dinucleotide phosphate, NAADP. Of these, the function of Ins(1,4,5)P_3 has been characterized first and in the most detail.

The concentration of free Ca^{2+} in the cytosol of "resting" cells is very low, about 10^{-7} M. One reason that the cell tries to keep the free Ca^{2+} concentration low is the ability of these ions to form poorly soluble complexes with inorganic phosphate. The low concentration of free cytosolic Ca^{2+} is opposed by a large storage capacity for Ca^{2+} in specific cytosolic compartments and by a high concentration in the extracellular region where Ca^{2+} is present at millimolar concentration. In the cytosol, Ca^{2+} is stored in the mitochondria and in special storage organelles of the endoplasmic reticulum. In the storage associated with the endoplasmic reticulum, Ca^{2+} exists in complex with the storage protein calreticulin. Calreticulin is a low-affinity Ca^{2+} binding protein with a high binding capacity. In the protein-bound and compartmentalized form, Ca^{2+} is not freely available but may be released in the process of signal transduction.

In muscle cells, Ca^{2+} is stored in the sarcoplasmic reticulum. The storage takes place particularly by binding to the storage protein calsequestrin. It is released from storage by a neural stimulus (see Section 6.5.1) and initiates muscle contraction.

From the "resting" state of low Ca^{2+}, cells can pass to the "on state" where Ca^{2+} increases and thereafter to the "off" state where the Ca^{2+} concentration is actively decreased.

The free Ca^{2+} concentration is subject to strict regulation, and targeted increase of Ca^{2+} is a universal means of controlling a vast array of metabolic and physiological reactions. Many processes are involved in Ca^{2+} regulation (Fig. 6.7), allowing the cell to shape Ca^{2+} signals in the dimensions of space, time and amplitude. Fig. 6.7 gives an overview of the main pathways leading to an increase or decrease of intracellular calcium.

6.5.1
Release of Ca^{2+} from Ca^{2+} Storage

Mobilization of Ca^{2+} from the Ca^{2+} stores of the endoplasmic reticulum takes place with the help of Ca^{2+} channels, of which two types stand out: the *$InsP_3$ receptors* and the *ryanodin receptors*. Both are ligand-gated Ca^{2+} channels, in which receptor and ion channel form a structural unit. The $InsP_3$ receptors and ryanodin receptors are localized in the endoplasmic and sarcoplasmic reticulum, respectively, and may be opened during the process of signal transduction (Fig. 6.8).

The $InsP_3$ Receptor

Binding of $InsP_3$ to the $InsP_3$ receptor leads to opening of the receptor channel, so that stored Ca^{2+} can flow into the cytosol. The $InsP_3$ receptor is a transmembrane protein, probably with two transmembrane domains in the vicinity of the C terminus. The active receptor is composed of four identical subunits. It is assumed that the Ca^{2+}

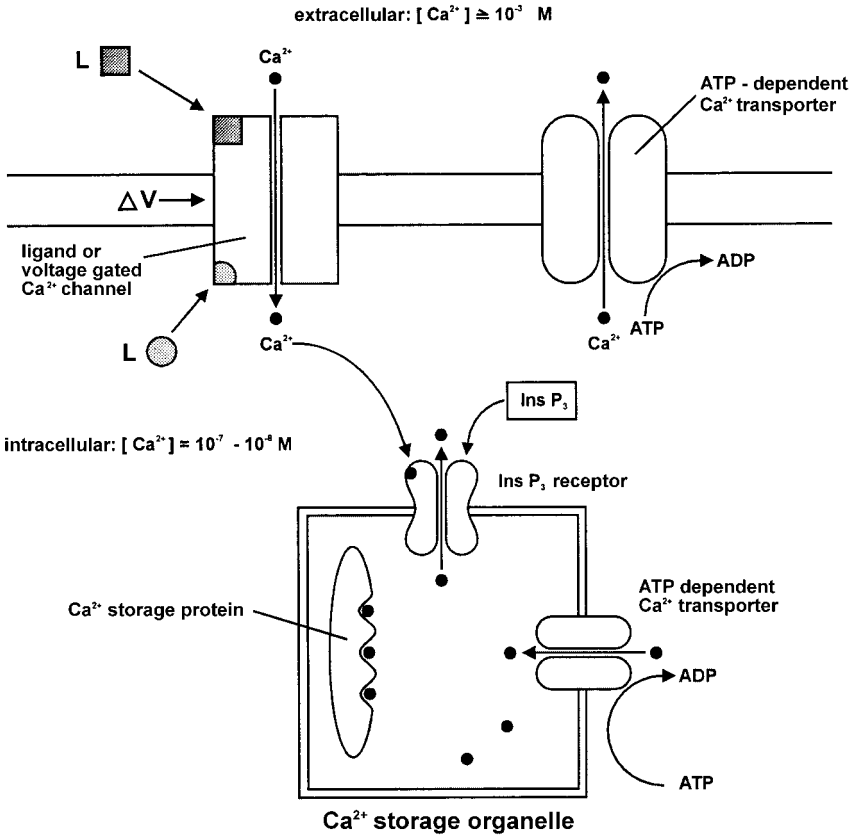


Fig. 6.7 Paths for increase and reduction of cytosolic Ca^{2+}concentration. Influx of Ca^{2+} from the extracellular space takes place via Ca^{2+}channels; the open state of these is controlled by binding of ligand L or by a change in the membrane potential (V). According to the type of ion channel, the ligand may bind from the cytosolic or the extracellular side to the ion channel protein. The entering Ca^{2+}binds to $InsP_3$ receptors on the membrane of Ca^{2+}storage organelles and induces, together with $InsP_3$, their opening. Ca^{2+}flows out of the storage organelle into the cytosol via the ion channel of the $InsP_3$ receptor. Transport of Ca^{2+} back into the storage organelles takes place with the help of ATP-dependent Ca^{2+} transporters.

channel is formed by the C-terminal transmembrane element and that the binding site for $InsP_3$ is localized in the large cytoplasmic region of the receptor. Opening of the $InsP_3$ receptor is subject to complex regulation involving Ca^{2+}, Mg^{2+} and ATP, in addition to $InsP_3$. Furthermore, $InsP_3$ receptors can be opened by interaction with proteins belonging to the group of Ca^{2+} binding proteins (review: Bootman, 2002). Regulation of the $InsP_3$ receptor by Ca^{2+} is of particular importance for the generation of Ca^{2+} waves. Low Ca^{2+} concentrations stimulate Ca^{2+} release by $InsP_3$ receptors, whereas high concentrations are inhibitory.

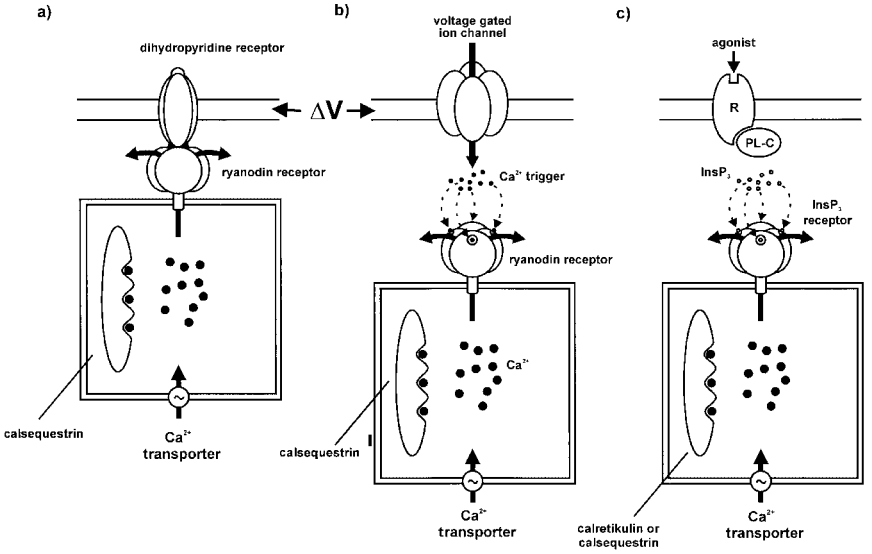


Fig. 6.8 Tetrameric Ca^{2+} channels and control of Ca^{2+} release. a) A change in the membrane potential (V) induces a conformational change in the dihydropyridine receptor of skeletal muscle; this is transmitted as a signal to the structurally coupled ryanodin receptor. Opening of the Ca^{2+} channel takes place and efflux of Ca^{2+} from the sarcoplasmic reticulum into the cytosol occurs. b) In cardiac muscle, the release of Ca^{2+} takes place by a Ca^{2+}-induced mechanism. A potential change V induces opening of voltage-gated Ca^{2+} channels. Ca^{2+} passes through, which serves as the trigger for release of Ca^{2+} from Ca^{2+} storage organelles by binding to ryanodin receptors on the surface of the storage organelles. c) Membrane-associated signaling pathways are activated by ligands and lead, via activated receptor and phospholipase C (PL-C) to formation of $InsP_3$ and to release of Ca^{2+} from storage organelles.

Ryanodin Receptor and Cyclic ADP Ribose

The ryanodin receptor takes its name from its stimulation by the plant alkaloid ryanodin. In all, it has a similar composition to the $InsP_3$ receptor and is involved in Ca^{2+} signal conduction in many excitatory cells (cells of banded and smooth musculature, neurons, etc.).

The open state of the ryanodin receptor is controlled in part by Ca^{2+}, which binds to the receptor and induces its opening. With the Ca^{2+}-induced opening of the ryanodin receptor, the cell has a cooperative, self-amplifying mechanism that can trigger a rapid increase in the Ca^{2+} concentration. An initial increase in Ca^{2+} concentration, induced by Ca^{2+} influx from the extracellular space due to opening of voltage-gated Ca^{2+} channels, for example, initiates the opening of ryanodin receptors. The additional Ca^{2+} emerging from the membrane compartments can now open more ryanodin receptors, leading to a steep increase in the Ca^{2+} concentration. As with the $InsP_3$ receptors, high Ca^{2+} concentrations inhibit Ca^{2+} flux through the channel.

In some cell types (including cardiac muscle cells and pancreatic cells), another "second messenger", the cyclic ADP-ribose (Fig. 6.9), is involved in opening the rya-

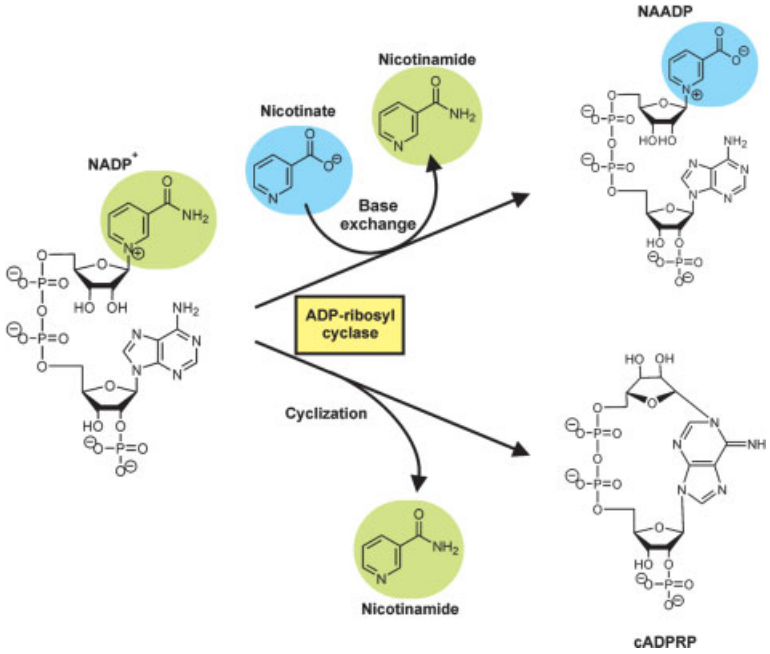


Fig. 6.9 Reactions of ADP ribosyl cyclase. Structures of NADP, nicotinic acid adenine dinucleotide phosphate (NAADP) and cyclic ADP-ribose phosphate (cADPRP). ADP-ribosyl cyclase, in base exchange mode, can catalyze replacement of the nicotinamide group of NADP (yellow) with nicotinic acid to generate NAADP. ADP-ribosyl cyclase can also catalyze cyclization of NADP to cADPRP.

nodin receptors . The cADP-ribose is formed from NADP by an enzymatic pathway with the help of an ADP-ribosyl cyclase.

Another messenger substance, nicotinic acid adenine dinucleotide phosphate, NAADP , can be also generated by the action of ADP-ribosyl cyclase (Fig. 6.9). NAADP has been discovered in brain and other tissues. NAADP releases Ca^{2+} independently of $InsP_3$ and cyclic ADP ribose signals from intracellular stores (review: Patel et al., 2001). The NAADP receptor however is not yet characterized.

A special coupling between extracellular Ca^{2+} influx and the ryanodin receptor exists in muscle cells. There, a voltage-dependent Ca^{2+} channel, the dihydropyridine receptor, is coupled directly to the cytoplasmic domain of the ryanodin receptor (see Fig. 6.8). A depolarization of the cell membrane is transmitted in this system via an electromechanical coupling directly to the gating state of the ryanodin receptor.

Tool Kit for Ca^{2+} Release

Overall, multiple pathways can be used for mobilising Ca^{2+} from the internal stores. (review: Bootman et al., 2001). A Ca^{2+} signaling 'toolkit' is available from which cells can select specific components to activate the internal Ca^{2+} stores and to generate a variety of different Ca^{2+} signals that suit their physiology. In summary, the following pathways can induce Ca^{2+} release from internal stores (Fig 6.10):

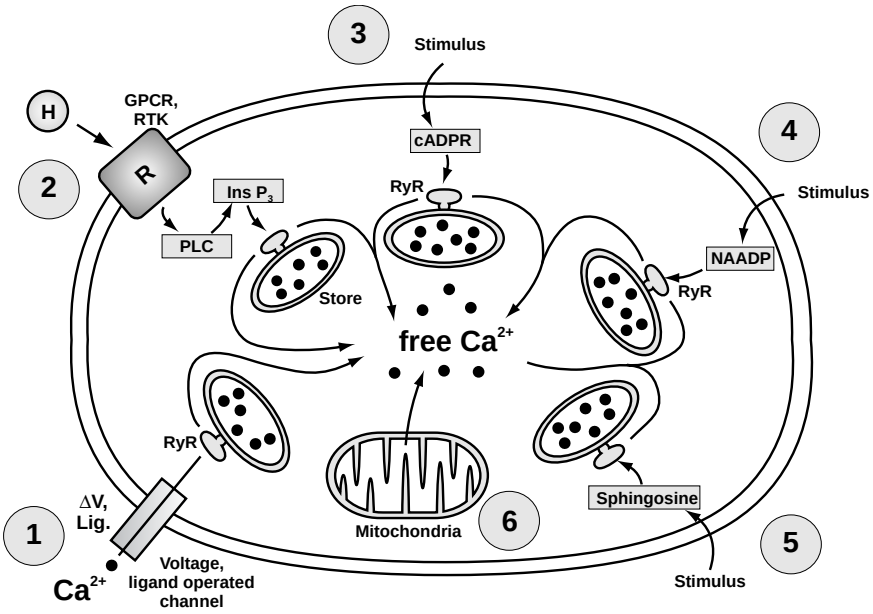


Fig. 6.10 Tools for Ca^{2+} release. The figure illustrates the major pathways for mobilising Ca^{2+} from internal stores. 1, Ca^{2+} induced Ca^{2+} release from ryanodine receptors (RyR) caused by the influx of Ca^{2+} through voltage- or ligand-gated channels on the outer cell membrane. This release may be also triggered by direct interaction of the channel with RYR. 2, PLC/InsP3 evoked release of Ca^{2+} from InsP3 receptors or ryanodine receptors. 3, cyclic ADP-ribose (cADPR)- evoked Ca^{2+} release. 4, nicotinic acid adenine dinucleotide phosphate (NAADP) evoked Ca^{2+} release. 5, Ca^{2+} release evoked by sphingosine. 6, Ca^{2+} release from mitochondria.

- Ca^{2+}-induced Ca^{2+} release from ryanodine receptors caused by influx of Ca^{2+} through voltage-operated Ca^{2+} channels on the plasma membrane.
- Cyclic ADP-ribose-evoked Ca^{2+} release.
- NAADP-evoked Ca^{2+} release
- InsP3-evoked Ca^{2+} release.
- Ca^{2+} release by interaction of $InsP_3$ receptors with calcium binding proteins
- Ca^{2+} release triggered by sphingolipids or leukotriene B4
- Ca^{2+} release from mitochondria

6.5.2 Influx of Ca^{2+} from the Extracellular Region

In the extracellular region, the Ca^{2+} concentration is over 10^{-3} M, which is very high in comparison to the free cytosolic Ca^{2+} concentration. The cell membrane contains a variety of different Ca^{2+} channel types that enable Ca^{2+} influx to take place from the extracellular region into the cytosol. One of the primary functions of Ca^{2+} entry is to charge up the internal stores, which can then release an internal Ca^{2+} signal.

The main Ca^{2+} influx channels are

- Voltage-gated channels are opened by a depolarization or change in membrane potential.
- Ligand-gated channels are activated by binding of an agonist to the exctracellular domain of the channel. Examples are provided by the acetylcholine receptor and the *N*-methyl-D-asparate receptor.
- Mechanically activated channels are present on many cell types and respond to mechanical stress.
- Store-operated channels are activated in response to depletion of the intracellular Ca^{2+} stores. The mechanism by which depletion of the internal stores is sensed by these channels is unknown.

In addition we know of Ca^{2+} channels that are controlled by G_α proteins (see Section 5.5.1) and Ca^{2+} channels that are gated by sphingolipids.

6.5.3 Removal and Storage of Ca^{2+}

The cytosolic Ca^{2+} concentration is generally only temporarily and is often only locally increased during stimulation of cells. The cell possesses efficient Ca^{2+} transport systems, which can rapidly transport Ca^{2+} back into the extracellular region or into the storage organelles. Ca^{2+}-ATPases, in particular, are involved in draining the cytosol of Ca^{2+} back into the extracellular region. The Ca^{2+}-ATPases perform active transport of Ca^{2+} against its concentration gradient, using the hydrolysis of ATP as an energy source. Other transport systems in the plasma membrane exchange Na^+ ions for Ca^{2+}. These *Na^+-Ca^{2+} exchange proteins* are located especially in muscle cells and in neurons. Ca^{2+}-ATPases, which can fill the empty Ca^{2+} storage, are also located in the membrane of the endoplasmic reticulum.

Opening of Ca^{2+} leads to a local increase in the cytosolic Ca^{2+} concentration from 10^{-7} M to 10^{-6} M. In this concentration region, the Ca^{2+} transport systems mentioned above work very efficiently. However, if an increase in Ca^{2+} concentration over 10^{-5} M takes place, e. g., because of cell damage, a level critical for the cell is reached. In this case, Ca^{2+} is pumped into the mitochondria with the help of Ca^{2+} transport systems localized in the inner membrane of the mitochondrion.

6.5.4 Temporal and Spatial Changes in Ca^{2+} Concentration

When the membrane channels and the intracellular release channels are activated, only brief pulses of Ca^{2+} are produced, since these channels have short open times. These "elementary" Ca^{2+} signals are localized around the channels and provide local control of many physiological reactions such as activation of other ion channels and nuclear-specific Ca^{2+} signals. The coordinated recruitment of many elementary Ca^{2+} release and entry channels allows the formation of global Ca^{2+} signals that persist

over a longer time and have a larger spatial distribution. Commonly, these global Ca^{2+} signals are of a pulsatile nature and appear as waves or spikes. The mechanisms by which waves and spikes are generated are diverse and are used in a cell-specific way. The differential sensitivity of the $InsP_3$ receptor and the ryanodin receptor to low and high Ca^{2+} concentrations is one mechanism that is assumed to contribute to Ca^{2+} wave and spike formation. A small increase in Ca^{2+} concentration due to an elementary Ca^{2+} signal will activate some release channels and allow the influx of more Ca^{2+} in a cooperative manner. The Ca^{2+} concentration increases until a threshold value is reached that is sufficient to inhibit Ca^{2+} influx through the channel, and the Ca^{2+} concentration falls again. In a further reaction, the released Ca^{2+} activates phospholipase C enzymes, leading to increased formation of $InsP_3$, which diffuses to $InsP_3$ receptors, bringing about release of more Ca^{2+}.

Various feedback mechanisms exist that ensure a decrease in Ca^{2+} concentration and concomitant peak and wave formation under conditions of constant exposure to stimulatory signals. One example is provided by the Ca^{2+}-dependence of subtypes of RGS proteins (Ishii et al., 2002) that can attenuate and shut down G protein-mediated signals. One of the effectors of G_q proteins is phospholipase Cβ, which stimulates Ca^{2+} release by $InsP_3$ formation. Activation of Gq proteins by G protein-coupled receptors can be terminated by RGS proteins, subtypes of which are Ca^{2+}/calmodulin dependent. These RGS proteins are stimulated under conditions of high Ca^{2+} and will therefore inhibit further activation of phospholipase Cβ and the induction of $InsP_3/Ca^{2+}$ signals.

Generally, the formation of cell-specific and highly variable global Ca^{2+} signals is based on the differential use of the various mechanisms that produce the Ca^{2+} "on" and "off" state. From the large Ca^{2+} signaling tool kit, each cell employs a specific set of channels and pumps to create signals that are highly variable in space and in time.

Ca^{2+} signals are very versatile signals that can store different information. Like electronic or optical signals in control engineering, the information content of Ca^{2+} signals may be determined by location, frequency, period and amplitude of the Ca^{2+} peak. Thus, the temporal sequence of Ca^{2+} signals has a regulatory function in many physiological processes. For example, in fluid secretion in the salivary glands of insects, an oscillatory Ca^{2+} signal is observed, the frequency of which is determined by the intensity of the external triggering signal.

How the frequency of an oscillating Ca^{2+} signal is decoded or integrated and incorporated into specific biochemical reactions is not understood. There is evidence that the CaM kinase II (see Section 7.5.2) is involved in decoding repetitive Ca^{2+} signals (De Koninck and Schulmann, 1998).

Amplitude modulation of Ca^{2+} signals is seen in B lymphocytes. These can create Ca^{2+} signals of higher or lower amplitude, initiated by the same extracellular signal, namely binding of an antigen to the B cell receptor. The different Ca^{2+} signals activate different effector molecules and trigger different intracellular reactions. Ca^{2+} signals of higher amplitude bring about activation of the transcription factor NFκB, whilst Ca^{2+} signals of lower amplitude lead to activation of transcription factor NF-AT (review: Berridge, 1997). Whether a Ca^{2+} signal of higher or lower amplitude is created depends on the differentiation state of the B lymphocytes.

6.6
Phosphatidyl Inositol Phosphates and PI3-Kinase

Several metabolic pathways lead from phosphatidyl inositol to compounds with "second messenger" character (review: Irvine and Schell, 2001). One main pathway, the formation of diacylglycerol and Ins(1,4,5)P_3 from PtdIns(4,5)P_2, has already been de-

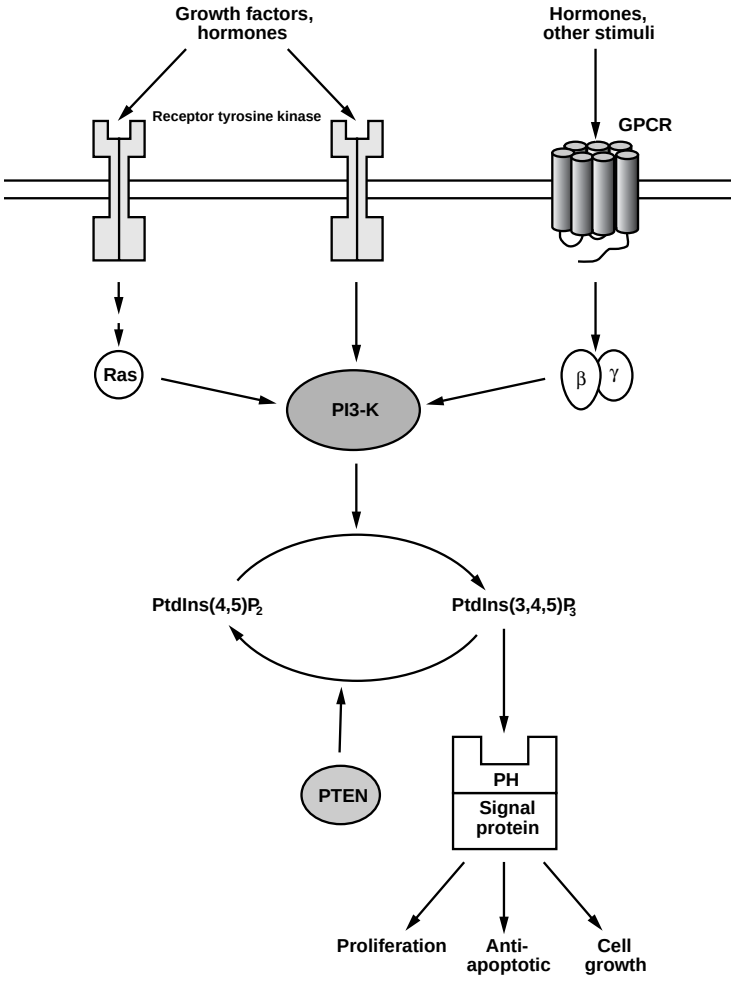


Fig. 6.11 Pathways of PI3-kinase activation. PI3-kinase can be activated by growth factor receptors, either by direct interaction or via the Ras protein. Another way of PI3-kinase activation uses the $\beta\gamma$-subunits of heterotrimeric G proteins liberated upon activation of G protein-coupled receptors, GPCR. The product of the PI3-kinase reaction is PtdIns(3,4,5)P_3 which binds to PH domains of various signaling proteins promoting their membrane association and activation. Overall, activation of PI3-kinase stimulates cell growth and proliferation and inhibits apoptosis. A suppressing effect is exerted by the tumor suppressor PTEN which hydrolyzes and thus inactivates PtdIns(3,4,5)P_3.

scribed in Section 6.4 and Fig. 6.4. Other major compounds of regulatory importance can be formed by phosphorylation at the 3' position of the inositol part of PtdIns. The reaction is catalyzed by a class of enzymes known as phosphatidyl inositide 3-kinases (PI3-kinases). The PI3-kinases phosphorylate various phosphatidyl inositol compounds at the 3' position. A major substrate is PtdIns(4,5)P_2, which is converted by PI3-kinase into PtdIns(3,4,5)P_3. This compound has an important function as an intracellular messenger. PtdIns(3,4,5)P_3 binds to PH domains of various signaling proteins promoting their membrane association (see Section 6.6.2). An overview of the function and regulation of PI3-kinase is given in Fig. 6.11.

6.6.1 PI3-Kinases

The family of PI3-kinases includes various enzymes that differ in the size of the catalytic subunit, the nature of the associated subunits and regulation (review: Krugmann and Welch, 1998; Anderson et al., 1999; Cantrell, 2001). Of the three classes (I – III) of PI3-kinases, only class I will be presented as an example here. Most members of class I are associated with a subunit that functions as an adaptor in signal transduction. The best investigated PI3-kinase, PI3-kinase α, is a heterodimer with adaptor function, made up of a catalytic subunit (p110α) and a regulatory subunit of 85 kDa (p85α). The p85α subunit has an SH3 domain, two SH2 domains and two Pro-rich domains. These domains function as binding modules, which the PI3-kinase uses for specific protein-protein interactions in the process of signal transduction and for association with other signal proteins (see Chapter 8).

Other members of class I of the PI3-kinases, such as PI3-kinase of the γ subtype, are stimulated by interaction with $\beta\gamma$-complexes (see Section 5.5.5) and have their own regulatory subunit. It is interesting that both a lipid kinase activity and a protein kinase activity have been identified in the catalytic domain of the PI3-kinase γ subtype in brain (Bondeva et al., 1998). Activation of the MAPK pathway (see Chapter 10) may take place via the protein kinase activity, so that this enzyme can produce a bifurcated signal: the lipid kinase activity stimulates the Akt kinase (see below), the protein kinase the MAPK pathway. Proliferation-promoting signals are transmitted via both pathways.

Most data are available for the p110α*p85α subtype of PI3-kinase. For brevity, this is referred to as PI3-kinase in the following. The PI3-kinase phosphorylates various PtdIns derivatives at the 3 position (see Fig. 6.4) *in vitro* (review Divecha and Irvine, 1995). A physiologically important substrate is PtdIns(4,5)P_2, which is converted to PtdIns(3,4,5)P_3 by PI3-kinase. PtdIns(3,4,5)P_3 is an intracellular messenger that has a regulatory effect in many elementary functions of the cell, such as growth control, chemotaxis and glycogen synthesis (see Section 6.6.2).

An important function in growth regulation is attributed to the PI3-kinase. PtdIns(3,4,5)P_3 is not detectable in resting cells. On stimulation of the cells with a growth factor, a rapid increase in PtdIns(3,4,5)P_3 occurs. An associated translocation of PI3-kinase to the membrane is observed.

Many observations indicate that PI3-kinase functions as a signal protein that receives signals on the cytoplasmic side of the cell membrane and transmits them further, although its primary role is to produce membrane-localized messenger substances.

PI3-kinase is activated via three pathways (see Fig. 6.11).

- **Interaction with activated receptor tyrosine kinases**
  The SH2 domain of the p85 subunit mediates an interaction with tyrosine residues on signal proteins involved in transduction of growth-regulating signals. Thus, binding of the PI3-kinase to tyrosine phosphate residues of the activated PDGF receptor is observed (see Section 8.1.4). Another binding partner is the insulin receptor substrate (IRS, see Section 8.5). In both cases, it is assumed that the binding of the SH2 domain of p85 to the tyrosine residue of the signal protein serves to target the PI3-kinase to its membrane-localized substrate. Furthermore, binding of p85 to phosphotyrosine residues of activated receptors appears to be accompanied by an allosteric activation of the catalytic subunit (Fig. 6.12).
  The interaction between PI3-kinase and the insulin receptor substrate links insulin signaling to the PI3-kinase pathways. Accordingly, most of the physiological functions of insulin are mediated by the PI3-kinase/Akt kinase pathway (review: Alessi and Downes, 1998).
- **Activation in the Ras pathway**
  The PI3-kinase has also been identified as a part of the Ras signaling pathway (see Chapter 9). Signals originating from transmembrane receptors can be transmitted from the Ras protein to PI3-kinase. In this case, the PI3-kinase acts as the effector molecule of the Ras protein.
- **Activation by the G$\beta\gamma$ dimer**
  G$\beta\gamma$ dimers directly activate the PI3-kinase β and γ subtypes. In this way, a variety of extracellular signals can be transmitted via G protein-coupled receptors (see Section 5.5.5) and G proteins to PI3-kinase and its effectors.

6.6.2 The Messenger Substance PtdIns(3,4,5)P_3

The products of the PI3-kinase reaction are different phosphoinositide derivatives phosphorylated at the 3 position, of which PtdIns(3,4,5)P_3 has the greatest regulatory importance. PtIns(3,4,5)P_3, like cAMP, has the function of a messenger substance that activates effector molecules in the sequence for further signal conduction. In contrast to cAMP, PtdIns(3,4,5)P_3 is localized in the cell membrane and performs its function in close association with processes at the cell membrane. The concentration of PtdIns(3,4,5)P_3 in the cell depends both on the rates of synthesis by PI3-kinases and the rates of hydrolysis of its phosphate residues. Several inositol polyphosphate phosphatases have been identified that remove the phosphates at position 3 or 5 of the inositol moiety. Among the inositol polyphosphate phosphatases with specificity for the 3-position, the PTEN phosphatase has been identified as a tumor supressor protein (see below).

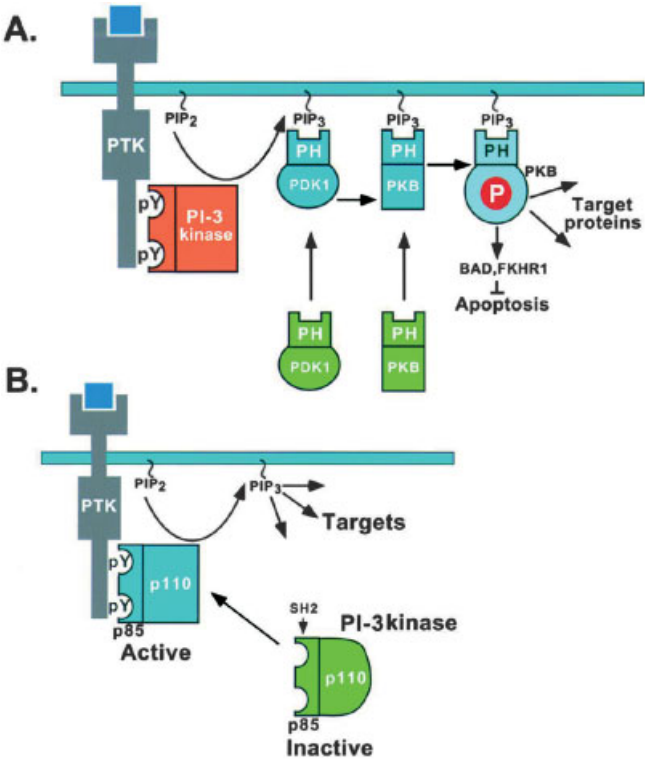


Fig. 6.12 (A) Activation of PKB (also known as Akt kinase) by membrane translocation.
PtdIns(3,4,5)P_3 generated in response to growth factor stimulation serves as a binding site for the PH domains of PDK1 and PKB. Membrane translocation is accompanied by release of an autoinhibition leading to activation of PDK1 and PKB kinase activities. Full activation of PKB requires phosphorylation by PDK1. Activated PKB phosphorylates a variety of target proteins that prevent apoptotic death (Bad) and regulate transcription (forkhead transcription factors, FKHR1) and other metabolic processes.
(B) Activation by a conformational change. Binding of the SH2 domains of p85, the regulatory subunit of PI-3 kinase to pTyr sites on activated receptors releases an autoinhibitory constraint that stimulates the catalytic domain (p110). PI-3 kinase catalyzes the phosphorylation of the 3' positions of the inositol ring of PtdIns(4)P and PtdIns(4,5)P2 to generate PtdIns(3,4)P_2 and PtdIns(3,4,5)P_3, respectively. (After Schlessinger, 2000).

The messenger function of PtdIns(3,4,5)P_3 is based on its specific binding to various protein modules on signaling proteins.

PtdIns(3,4,5)P_3 exerts most of its cellular functions by binding to pleckstrin homology domains (PH domains) of signal proteins. PH domains are found as independent protein modules in many signal proteins (see Section 8.2.4) that mediate protein-lipid and possibly also protein-protein interactions. PtdIns(3,4,5)P_3 formed by PI3-kinase serves to recruit PH domain-containing proteins to the membrane and to involve them

in signal conduction. In addition to membrane targeting, PtdIns(3,4,5)P_3 binding to PH domains can also bring about an allosteric activation of effector proteins.

Specific binding of PtdIns(3,4,5)P_3 has also been reported for other protein modules found in signaling proteins, namely SH2 domains, PTB domains, the C2 domain of protein kinase C and the FYVE ring finger domain of several membrane proteins.

6.6.3 Akt Kinase and PtdIns(3,4,5)P_3 Signaling

PtdIns(3,4,5)P_3 formed by PI3-kinase regulates the activity of a series of protein kinases, including the Ser/Thr-specific Akt kinases, protein kinase C enzymes (see Section 7.4), and the tyrosine-specific Tec kinases. Only the regulation of Akt kinase will be discussed in the following.

The first target protein of PtdIns(3,4,5)P_3 to be characterized was Akt kinase, also known as protein kinase B (PKB). Akt kinase is a Ser/Thr-specfic protein kinase which regulates multiple biological processes including glucose metabolism, apoptosis, gene expression, and cellular proliferation. The signaling pathway for Akt kinase shown in Fig. 6.12 illustrates the central role of PI3-kinase and PtdIns(3,4,5)P_3 in growth factor-controlled signal paths that lead from the cell membrane into the cytosol and the nucleus.

In the Akt signaling pathway (review: Chan et al., 1999, Testa and Bellacosa, 2001), first an extracellular growth factor activates the corresponding transmembrane receptor (e. g., PGDF receptor, see Section 8.1). Consequently, tyrosine phosphorylation takes place on the cytoplasmic domain of the receptor. The tyrosine residues serve as docking sites for the SH2 domain of the p85 subunit of the PI3-kinase. The associated translocation of PI3-kinase is synonymous with its activation. The PtdIns(3,4,5)P_3 formed binds to the PH domain of the signal protein next in sequence, the Akt kinase, inducing its activation.

Activation of Akt kinase proceeds in a multi-step process with regulation by PtdIns(3,4,5)P_3 as the critical step. PtdIns(3,4,5)P_3 binds to the PH domain of Akt kinase and thereby recruits the enzyme to the cell membrane. This step is thought to be a prerequisite for a subsequent phosphorylation of a Thr residue in the activation loop (see Chapter 7 on protein kinases) and an N-terminal Ser residue of Akt. The protein kinase responsible for this step is named phosphoinositide-dependent protein kinase 1 (PDK1, review: Vanhaesebroeck and Alessi, 2000), and this also contains a PH domain with high affinity for PtdIns(3,4,5)P_3. Membrane targeting and the double phosphorylation activate Akt, allowing the phosphorylation of downstream substrates like the Bad protein (see Section 15.7.1), pro-caspase 9, the transription factor CREB (see Section 1.4.5.2), glycogen synthase kinase (GSK), and 6-phosphofructo-2-kinase.

According to this mechanism, Akt kinase regulates central metabolic pathways of the cell. It is a major regulator in insulin-dependent metabolic pathways. The PI3-kinase/Akt pathway is activated by insulin and thereby mediates many of the metabolic effects of insulin, including glucose transport, lipid metabolism, glycogen synthesis and protein synthesis (see Section 1.5.6). Furthermore, by phosphorylating pro-

teins of the apoptotic machinery, Akt has an inhibitory influence on programmed cell death, apoptosis (see Section 15.7.1). Furthermore, Akt has been shown to have a promoting influence on cell division. In accordance with a critical role in cell proliferation, Akt kinase has been identified as a classical viral oncogene.

The great importance of PtdIns(3,4,5)P_3 metabolism for growth regulation is illustrated by the observation that an enzyme of PtdIns(3,4,5)P_3 metabolism has been identified as a tumor suppressor protein (Wu et al., 1998). PTEN tumor suppressor protein has lipid phosphatase activity that is specific for hydrolysis of PtdIns(3,4,5)P_3. It is assumed that PTEN lipid phosphatase is a negative regulator of the Akt pathway, acting by lowering the concentration of PtdIns(3,4,5)P_3 and counteracting stimulation of Akt kinase (see Fig. 6.11). Because of the strong cell proliferation-promoting and antiapoptotic activity of the Akt kinase pathway, lowered concentrations of PtdIns(3,4,5)P_3 will have an antiproliferative and proapoptotic effect and will thus inhibit tumor formation, explaining the tumor-suppressing activity of PTEN. In accordance with this model, a functional inactivation of the PTEN phosphatase is observed in many tumors.

6.6.4
Functions of PtIns(4,5)P_2

Inositol phosphatides have another role in the formation of microfilaments of the cytoskeleton.

Polymerization and depolymerization of actin, the main component of microfilaments, is controlled by a series of proteins, the activity of which is regulated by Ca^{2+} and/or PtdIns(4,5)P_2. The Ca^{2+}-regulated proteins (see Section 6.7) are chiefly involved in processes of depolymerization of actin. Many of the proteins involved in the opposite process, actin polymerization, have specific binding sites for PtdIns(4,5)P_2 and PtIns(4)P and are regulated by the availability of phosphoinositides. Examples of such proteins are profilin, gelsolin, villin and talin.

6.7
Ca^{2+} as a Signal Molecule

Ca^{2+} is a central signal molecule of the cell. Following a hormonal or electrical stimulation, an increase in cytosolic Ca^{2+} occurs, leading to initiation of other reactions in the cell. As outlined above, this increase is limited in time and in space and allows the formation of a variety of differently shaped Ca^{2+} signals. Examples of Ca^{2+}-dependent reactions are numerous and affect many important processes of the organism, including

- muscle contraction
- vision process
- cell proliferation
- secretion
- cell motility, formation of the cytoskeleton

– gene expression
– reactions of intermediary metabolism.

Ca^{2+} signals in the form of temporally and spatially variable changes in Ca^{2+} concentration serve as elements of intracellular signal conduction in many signaling pathways.

Three main paths for increase in Ca^{2+} concentration stand out (Table 6.1; Fig. 6.5; Fig. 6.7):
– G-protein-mediated signaling pathways
– signaling pathways involving receptor tyrosine kinases
– influx of Ca^{2+} via voltage- or ligand-gated Ca^{2+} channels.

What is the Basis of the Function of Ca^{2+} as a Signal Molecule?

The information encoded in transient Ca^{2+} signals is deciphered by various intracellular Ca^{2+}-binding proteins that convert the signal into a wide variety of biochemical changes. There are two principle mechanisms by which Ca^{2+} can perform a regulatory function:

- **Direct activation of proteins**
  Many proteins have a specific binding site for Ca^{2+}, and their activity is directly dependent on Ca^{2+} binding. The available Ca^{2+} concentration thus directly controls the activity of these proteins (see Table 6.2).
  There are many enzymes that have a specific Ca^{2+}-binding site in the active center and for which Ca^{2+} has an essential role in catalysis. An example of a Ca^{2+}-dependent enzyme is phospholipase A2. Phospholipase A2 catalyzes the hydrolysis of fatty acid esters at the 2' position of phospholipids (see Fig. 5.28), where Ca^{2+} plays

Tab. 6.1 Receptors of the plasma membrane that mediate increase of intracellular Ca^{2+}.

mediated via Phospholipase Cβ	mediated via Phospholipase Cγ	direct
α_1 adrenergie receptor	epidermal growth factor-receptor	nicotinic acetylcholine receptor
muscarinic acetylcholin receptors	platelet derived growth factor receptor	glutamate receptors
glucagon	Fibroblast growth factor receptor	
serotonin receptor	T-cell receptor	
vasopressin receptor		
ocytocin receptor		
angiotensin II receptor		
thrombin receptor		
bombesin receptor		
bradykinin receptor		
tachykinin receptor		
thromboxan receptor		

Tab. 6.2 Ca^{2+} binding proteins.

protein	function
troponin C	modulator of muscle contraction
caldesmon	modulator of muscle contraction
α-actinin	bundling of actin
villin	organization of actin filaments
calmodulin	modulator of protein kinases and other enzymes
calcineurin B	protein phosphatase
caplain	protease
phospholipase A2	release of arachidonic acid
proteinkinase C	ubiquitous protein kinase
Ca^{2+}-activated K^+ channel	effektor of hyperpolarization
$InsP_3$ receptor	intracellular Ca^{2+} release
ryanodin receptor	intracellular Ca^{2+} release
Na^+/Ca^{2+} transporter	exchange of Na^+ and Ca^{2+} via the cell membrane
Ca^{2+} ATPase	transport of Ca^{2+} through cell membrane
recoverin	regulation der guanylyl cyclase
parvalbin	Ca^{2+} storage
calreticullin	Ca^{2+} storage
calbindin	Ca^{2+} storage
calsequestrin	Ca^{2+} storage

an essential role. The enzyme has two Ca^{2+} ions bound tightly at the active center. One of the two Ca^{2+} ions is directly involved in catalysis. It binds the substrate in the ground state and also helps to neutralize charge in the transition state of ester hydrolysis. The second Ca^{2+} ion is assigned a role in the stabilization of the transition state, in addition to a structural function.

Another example of a Ca^{2+}-regulated enzyme is protein kinase C (see Section 7.4). We also know of many proteins without enzyme activity that have Ca^{2+}-regulated functions. Proteins involved in the complex process of polymerization and depolymerization of the cytoskeleton are also often regulated by Ca^{2+} binding . These include the annexins, fimbrin, gelsolin and villin. The latter two are also regulated via $PtInsP_2$. Ca^{2+} and $PtInsP_2$ also have antagonistic effects on the polymerization state of microfilaments. Ca^{2+} promotes depolymerization of microfilaments and $PtInsP_2$ promotes their polymerization.

- **Binding to Ca^{2+} receptors**

  Another central mechanism of signal transduction via Ca^{2+} is its binding to Ca^{2+}-binding proteins also known as *Ca^{2+} receptors*. The receptor proteins function as regulatory proteins that couple the Ca^{2+} signal to other signaling proteins. The Ca^{2+} receptors are Ca^{2+} sensors that activate target proteins in response to changes in Ca^{2+} concentration. Increases in Ca^{2+} above the concentration of the resting state (ca. 10^{-7} M) lead to specific binding of Ca^{2+} to Ca^{2+}-binding sites on the receptor and

concomitant conformational changes that modulate the interaction with downstream target proteins.

We know of many Ca^{2+} receptors that require an increase in Ca^{2+} concentration for activation of a target protein. There are also Ca^{2+} receptors with an activating function inhibited by high Ca^{2+} concentrations, so that they are only active at low Ca^{2+} concentrations.

6.7.1 Calmodulin as a Ca^{2+} Receptor

The most widespread Ca^{2+} receptor is *calmodulin*. Calmodulin is a small protein of ca. 150 amino acids (review: Hoeflich and Ikura, 2002). The structure of the Ca^{2+}/calmodulin complex has two globular domains that are separated by a long α-helical section (Fig. 6.13a). Both globular domains have two binding sites for Ca^{2+}. Ca^{2+} is bound via a characteristic helix-loop-helix structure, also known as an EF structure. Similar EF structures are found in many, but not all, Ca^{2+}-binding proteins. The binding affinity

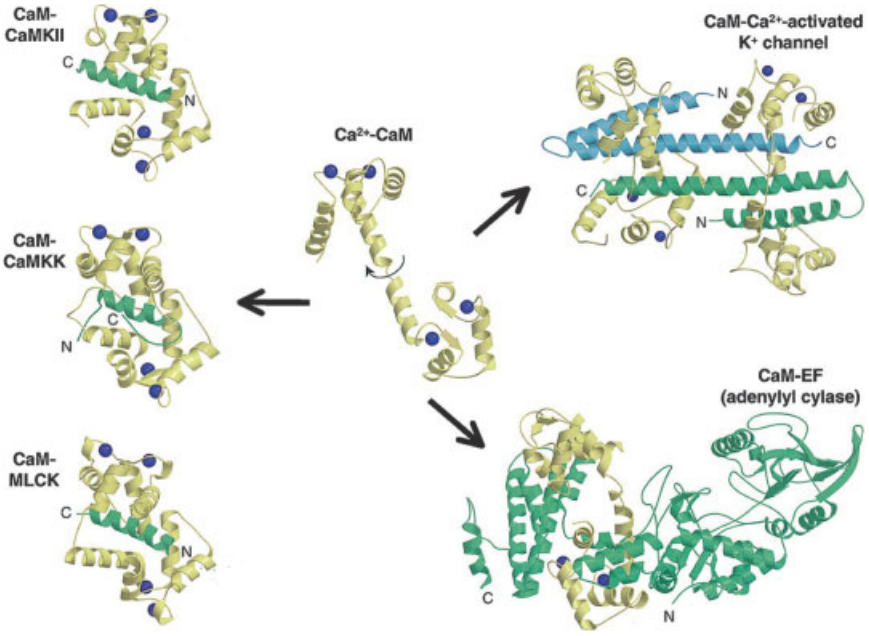


Fig. 6.13 Comparison of different Ca^{2+}/Calmodulin structures (from Hoeflich and Ikura, 2002). The figure illustrates the different conformations of calmodulin when bound to target protein kinases. Calmodulin is shown in yellow and calcium ions are depicted in blue. The interaction with the calmodulin binding domain of the protein kinases is mediated by short helices shown in green and blue. CaM-CaMKII: Ca^{2+}/calmodulin-dependent protein kinase II; CaM-CaMKK: Ca^{2+}/calmodulin-dependent protein kinase kinase; CaM-MLCK: Ca^{2+}/calmodulin-dependent myosin light chain kinase; CaM-EF: Ca^{2+}/calmodulin-dependent edema factor, an adenylyl cyclase, from *Bacillus anthracis*.

for Ca^{2+} varies widely for different EF structures. The dissociation constants of Ca^{2+} binding to EF-hands lie between 10^{-5} M and 10^{-9} M.

Calmodulin comprises four EF-hands organized in two globular domains. It binds 4 Ca^{2+} ions with an affinity ($K_D = 5 \times 10^{-7}$ M to 5×10^{-6} M) that fits into the intracellular Ca^{2+} concentrations exhibited by most cells. The degree of Ca^{2+} binding is therefore well suited to mirror changes in Ca^{2+} during Ca^{2+} signaling. Significant structural changes are induced in calmodulin upon Ca^{2+} binding leading to the exposure of hydrophobic residues in each of the two domains. The conformational change is akin to a Ca^{2+}-controlled unfolding of calmodulin, and it is assumed that interactions with the target proteins of Ca^{2+}/calmodulin are mediated by the newly exposed hydrophobic residues.

With the help of NMR measurement, it has been shown that the Ca^{2+}/calmodulin complex has a flexible structure. Flexibility is probably of great importance for the function of Ca^{2+}/calmodulin. Structural information on Ca^{2+}/calmodulin bound to substrates is only available for peptides derived from target proteins. In the complex with peptide substrates (Fig. 6.13), Ca^{2+}/calmodulin has a collapsed structure in which the two globular domains are much closer together than in free Ca^{2+}/calmodulin, and it wraps around and sequesters the helical calmodulin-binding peptides.

The mechanisms by which calmodulin regulates its target proteins are diverse and can be categorized into several classes. The most important ones are (review Chin and Means, 2000)

- Irreversible binding of calmodulin to the target protein irrespective of Ca^{2+}. One example is phosphorylase kinase, an enzyme that contains calmodulin as a firmly bound subunit and is activated in the presence of Ca^{2+}.
- Formation of inactive, low-affinity complexes with calmodulin at low Ca^{2+} concentrations and transition to an active complex in the presence of high Ca^{2+}. This class includes the protein phosphatase calcineurin.
- Activation by Ca^{2+}/calmodulin. Target proteins exhibiting this more conventional behavior include the Ca^{2+}/calmodulin-dependent protein kinases (see Section 7.5).
- Inhibition by Ca^{2+}/calmodulin. This class includes members of the G-protein-coupled receptor kinases and subtypes of the InsP3 receptor.

From the structures of the substrates and their complexes with Ca^{2+}/calmodulin (review Hoeflich and Ikura, 2002), two main mechanisms of substrate activation have emerged (Fig. 6.14). By one mechanism an autoinhibitory element is displaced from the active site of the target enzyme relieving autoinhibition. Another protein activation mechanism of Ca^{2+}/calmodulin uses a remodeling of the active site of the target protein.

6.7.2
Target Proteins of Ca^{2+}/Calmodulin

The Ca^{2+}/calmodulin complex is a signal molecule that is involved in many signal transduction pathways. Ca^{2+}/calmodulin is involved, e. g., in regulation of proliferation, mitosis, neuronal signal transduction, muscle contraction and glucose metabo-

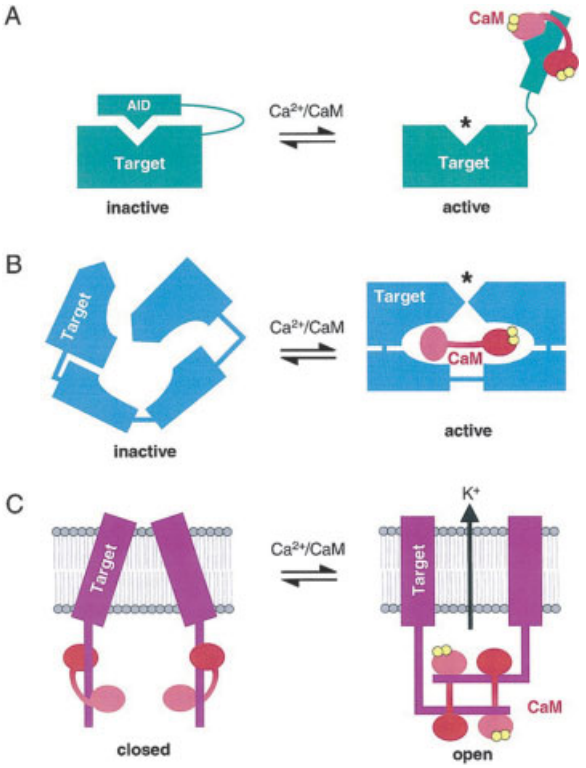


Fig. 6.14 Mechanisms of activation of target proteins by Ca^{2+}/calmodulin (after Hoeflich and Ikura, 2002).
A) Binding of Ca^{2+}/calmodulin relieves autoinhibition (CaM-kinases, calcineurin).
B) Ca^{2+}/calmodulin remodels the active site inducing an active conformation (anthrax adenylyl cyclase).
C) Ca^{2+}/calmodulin-induced dimerization of K^+-channels.
AID: autoinhibition domain.

lism. Different calmodulin subtypes are known which regulate different target proteins. The best characterized target proteins are the calmodulin-dependent adenylyl cyclases, phosphodiesterases, the protein phosphatase calcineurin (see Section 7.6.5), protein kinases like the CaM kinases (see Section 7.5), and the myosin light chain kinase (MCLK), involved in contraction of smooth musculature.

6.7.3 Other Ca^{2+} Receptors

The cell contains other Ca^{2+} receptors, some of which are related to calmodulin, which occur in specialized tissue and perform specific functions there.

Troponin C in muscle is structurally closely related to calmodulin. It has 4 EF structures, of which only two can be occupied by Ca^{2+}. Troponin C is a component of the contraction apparatus of muscle. Ca^{2+} binding to troponin C induces a conformational change in the troponin complex that leads to contraction of muscle.

Another regulatory Ca^{2+} receptor is *recoverin*, which performs an important control function in the signal transduction cascade of the vision process, by inhibiting the activity of rhodopsin kinase (see Section 5.3.4).

Recoverin is a Ca^{2+} receptor with four EF structures and two Ca^{2+}-binding sites; it can exist in the cytosol or associated with the membrane and has an N-terminal myr-

istoyl residue as a lipid anchor. The distribution between free and membrane-associated forms is regulated by Ca^{2+}. Binding of Ca^{2+} to recoverin leads to its translocation from the cytosol to the membrane of the rod cells. Structural determination of recoverin in the Ca^{2+}-bound and Ca^{2+}-free forms (Ames et al., 1997) indicates that membrane association of recoverin is regulated by a *Ca^{2+}-myristoyl switch* (see Section 3.7.5). The myristoyl residue can adopt two alternative positions in recoverin. In the absence of Ca^{2+}, recoverin exists in a conformation in which the myristoyl residue is hidden in the inner of the protein and is not available for membrane association. On Ca^{2+} binding, a conformation change of recoverin takes place; the myristoyl residue moves to the outside and can now associate with the membrane.

6.8 Diacylglycerol as a Signal Molecule

During cleavage of $PtInsP_2$ by phospholipase C, two signal molecules are formed, $InsP_3$ and diacylglycerol. Whilst $InsP_3$ acts as a diffusible signal molecule in the cytosol after cleavage, the hydrophobic diacylglycerol remains in the membrane. Diacylglycerol can be produced by different pathways, and it has at least two functions (Fig. 6.15). Diacylglycerol is an important source for the release of arachidonic acid, from which biosynthesis of prostaglandins takes place. The glycerine portion of the inositol phosphatide is often esterifed in the 2' position with arachidonic acid; arachidonic acid is cleaved off by the action of phospholipases of type A2.

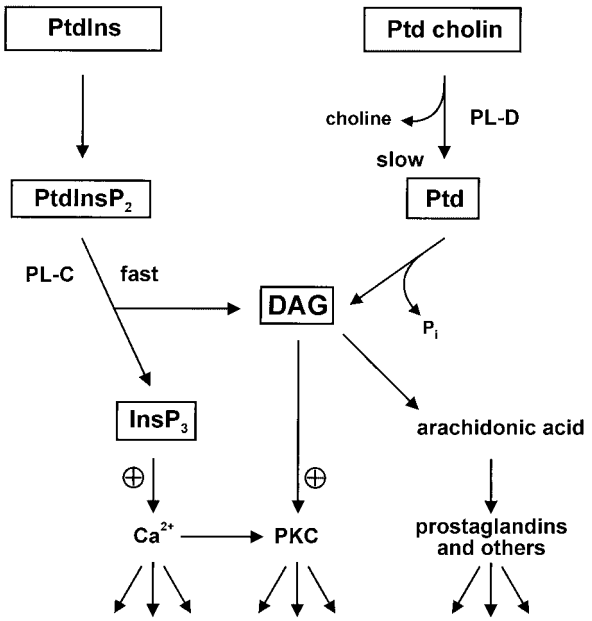


Fig. 6.15 Formation and function of diacylglycerol. The figure schematically shows two main pathways for formation of diacylglycerol (DAG). DAG can be formed from $PtdInsP_2$ by the action of phospholipase C (PL-C). Another pathway starts from phosphatidyl choline. Phospholipase D (PL-D) converts phosphatidyl choline to phosphatidic acid (Ptd), and the action of phosphatases results in DAG. Arachidonic acid, the starting point of biosynthesis of prostaglandins and other intracellular and extracellular messenger substances, can be cleaved from DAG. PKC: protein kinase C; PtdIns: phosphatidyl inositol.

The second important regulatory function of diacylglycerol is stimulation of protein kinase C (see Section 7.4). Protein kinase C is a protein kinase occurring in almost all cells and has a regulating effect on many reactions of the cell. Characteristic for protein kinase C is its stimulation by Ca^{2+}, diacylglycerol and phosphatidyl serine.

6.9 Other Lipid Messengers

In addition to the membrane-associated messenger substances diacylglycerol and PtdIns(3,4,5)P_3 mentioned above, other lipophilic compounds have been identified that are specifically formed in the process of signal transduction and which function as messenger substances. Two such compounds are presented below.

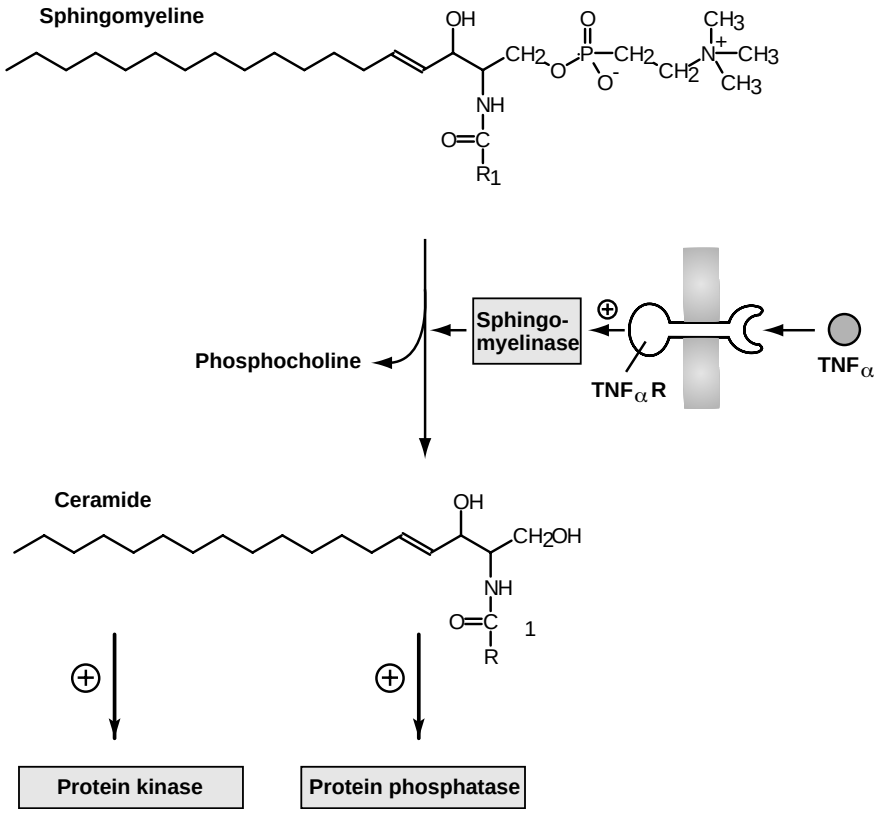


Fig. 6.16 Formation and function of the messenger substance ceramide. The starting point for the synthesis of ceramide is sphingomyelin, which is converted to phosphocholine and ceramide by the action of a sphingomyelinase. Sphingomyelinase is activated via a pathway starting from tumor necrosis factor α (TNFα) and its receptor. Ceramide serves as an activator of protein kinases and protein phosphatases. R1: fatty acid side chain.

- **Ceramide**
  Ceramide is a lipophilic messenger that regulates diverse signaling pathways involving apoptosis, stress response, cell senescence, and differentiation. For the most part, ceramide's effects are antagonistic to cell growth and survival (review: Ruvolo, 2001). The starting point for the formation of ceramide is sphingomyelin, which occurs especially in the outer layer of the plasma membrane. Ceramide is produced from sphingomyelin by the action of the enzyme sphingomyelinase (Fig. 6.16). Sphingomyelinase has similar cleavage specificity to phospholipase C, in that it cleaves an alcohol-phosphate bond. Activation of sphingomyelinase is observed in response to diverse stress challenges including irradiation, exposure to DNA-damaging agents or treatment with pro-apoptotic ligands like tumor necrosis factor α (TNFα, Chapter 11 and 15). Because of these properties, ceramide is a potent apoptogenic agent.
  The ceramide (or also ceramide-1-phosphate), produced by the action of sphingomyelinase, is a membrane-located messenger substance that binds to and activates various downstream targets including stress-activated protein kinases like the c-Jun-terminal protein kinase (JNK, see Section 10.2.2) and the protein phosphatases 1 and 2 (see Section 7.6). The activation of the phosphatases appears to downregulate critical pro-growth-signaling molecules like protein kinase C subtypes and the tumor suppressor protein pRb (see Section 13.4.2). From its location and synthesis, ceramide may be compared to diacylglycerol; however they have opposite effects on cell growth. While diacylglycerol stimulates cell growth via protein kinase C, ceramide is a potent inhibitor of cell proliferation.
- **Lysophosphatidic acid, LPA**
  Messenger substances derived from phospholipids can also function as hormones and serve for communication between cells. An important *extracellular* messenger substance formed from phospholipids is *lysophosphatidic acid* (1-acyl-sn-glycerine-3-phosphate). Lysophosphatidic acid (lysophosphatidic acid, LPA) is released by platelets and other cells and reaches its target cells via the circulation. As a product of the blood clotting process, LPA is an abundant constituent of serum, where it is found in an albumin-bound form.
  LPA binds and activates specific G protein-coupled receptors found in many cells (review: Kranenburg and Moolenaar, 2001). The LPA receptor can transmit the signal to G_q-, G_i- or G_{12}-proteins. If G_q is involved, an $InsP_3$ and Ca^{2+} signal is produced in the cell, whereas signal conduction via G_i- or G_{12}-proteins flows into the Ras pathway or activates the Rho proteins, respectively (see Chapter 9).

6.10 The NO Signaling Molecule

The biological importance of nitrogen monoxide (NO) as a messenger substance was originally recognized in connection with contraction and relaxation of blood vessels. In the meantime, it has become clear that NO is a universal messenger substance that is found in nearly all living cells. NO takes part in intercellular and intracellular com-

munication in higher and lower eucaryotes and it is also found in bacteria and in plant cells.

NO fulfills many criteria required to qualify as an intracellular and intercellular messenger. It is formed with the help of specific enzyme systems activated by extracellular and intracellular signals. It is synthesized intracellularly and reaches its effector molecules, which may be localized in the same cell or in neighboring cells, by diffusion. Thus, it has the character of an autocrine or paracrine hormone, as well as an intracellular messenger.

Classical extracellular messengers, such as the steroid hormones, bind to their corresponding receptor but do not undergo any chemical reaction with the latter, and the binding event is sufficient to activate the receptor. NO can also accumulate at the target protein; however, it can – in a clear difference from classical messengers – bring about a *covalent modification* at the target protein. The modification of the target protein is, for the most part, reversible and transient, and the modified target protein can transmit the signal to other effector proteins. Signal transduction via NO mostly takes place by covalent modification of the target protein, in which *redox reactions* are involved in passing on the signal.

6.10.1 Reactivity and Stability of NO

NO is a radical that is water soluble and can cross membranes fairly freely by diffusion. Because of its radical nature, NO has a lifetime in aqueous solution of only ca. 4 s. Important reaction partners of NO in biological systems are oxygen O_2, the O_2^- radical and transition metals in free or complex form, e. g. Fe^{2+} in heme (review: Stamler et al., 2001). Furthermore, NO readily reacts with nucleophilic centers in peptides and proteins, in particular with the SH groups of Cys residues (Fig. 6.17).

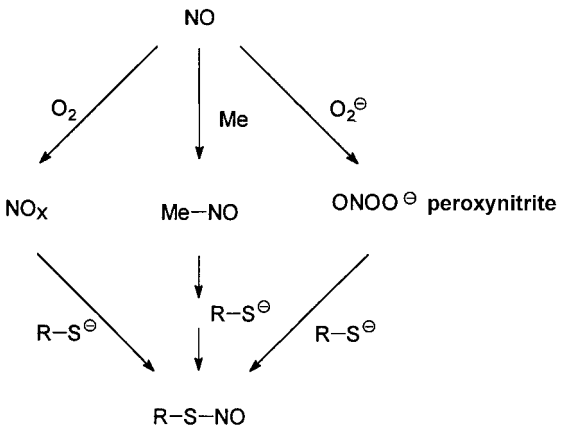


Fig. 6.17 Reactions of NO in biological systems. NO reacts in biological systems primarily with O_2, with the superoxide anion O_2^- and with transition metals (Me). The products of the reaction, -NO_x, metal-NO adducts (Me-NO) and peroxynitrite ($OONO^-$) react further by nitrosylation of nucleophilic centers. In the cell, these are especially –SH (or thiolate-S^-) groups of peptides and proteins (RS^-).

Binding to Metal Ions

Physiologically important reactions of NO with metals take place at metal ion centers of enzymes and at Fe ions of heme proteins. Binding of NO can lead to oxidation of the complexed metal ions.

Reaction with Thiolene: S-Nitrosylation

Reaction of NO with cysteine residues requires firstly the one-electron oxidation of NO to the NO^+ ion. This then accumulates at thiolate groups of peptides or enzymes (RS^-), forming S-nitrosyl groups RS-NO. The mechanism of S-nitrosylation of proteins by NO in the cell is not clear. From the S-nitrosyl group, the NO^+ can be cleaved off in a heterolytic reaction and transferred to nucleophilic centers (thiol groups) in acceptor proteins, creating a redox signaling chain. A crucial aspect of S-nitrosylation of peptides or proteins is the increased stability of the –S–NO grouping in comparison to free NO. Whilst free NO is only stable for a few seconds, the –S–NO grouping, e. g., in glutathione, has a half life of several hours.

6.10.2 Synthesis of NO

NO is formed enzymatically from arginine with the help of *NO synthase*, producing citrulline (Fig. 6.18). Citrulline and arginine are intermediates of the urea cycle, and arginine can be regenerated from citrulline by urea cycle enzymes.

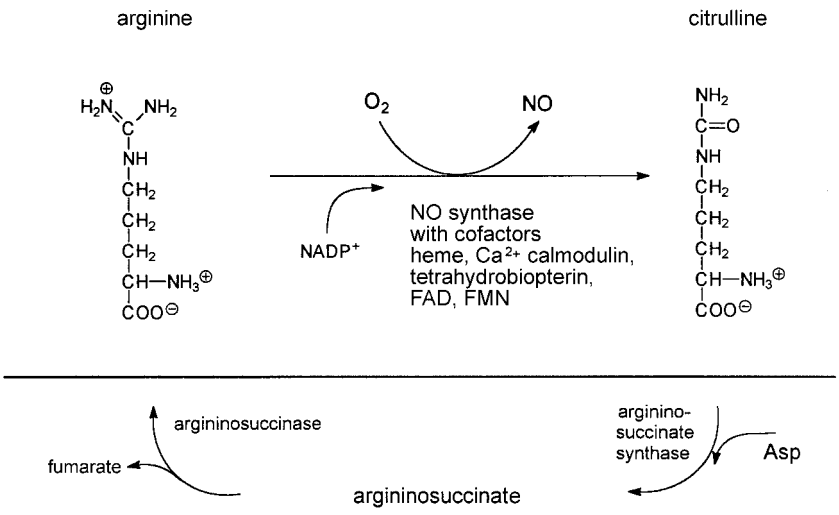


Fig. 6.18 Biosynthesis of NO. The starting point of NO synthesis is arginine. Arginine is converted by NO synthase, together with O_2 and NADP, to NO and citrulline. Arginine can be regenerated from citrulline via reactions of the urea cycle.

There are three forms of NO synthase (NOS I, NOS II, and NOS III), and these have differing sensitivity to Ca^{2+} (review: Alderton et al., 2001). NO synthases of types I and III require Ca^{2+} for activity and are predominantly consitutively activated. The activity of NOS II, in contrast, is not regulated by Ca^{2+}, and the enzyme is inducible at the level of the gene.

The NO synthases are enzymes of complex composition (MW ca. 300 kDa) that are active as dimers but can also exist as inactive monomers. Furthermore, the NO synthases of types I and III undergo complex regulation by Ca^{2+}/calmodulin. The following cofactors and substrates are required for reaction of the NO synthase:

- FAD, FMN
- L-arginine
- tetrahydrobiopterin
- heme
- NADPH
- O_2.

6.10.3 Physiological Functions and Attack Points of NO

The physiological importance of NO is due to both its regulatory and its toxic functions. NO produced in a regulated manner serves to control a multitude of cellular reactions. However, excessive production of NO can have toxic effects and is also used in mammals for antimicrobial action (review: Hausladen and Stamler, 1999; Toreilles, 2001, Stamler et al., 2001).

Toxic Action of NO and Nitrosative Stress

When NO is produced in excess amounts and in a less than regulated fashion, nonspecific reactions with various cell constituents including proteins, lipids and DNA are observed. This situation has been termed nitrosative stress in analogy to oxidative stress caused by the generation of reactive oxygen species, ROS. Nitration, nitrosation and oxidation of proteins, lipids and DNA can occur under these conditions and can lead to damage of cellular functions and eventually to cell death. Especially when excess NO and ROS are produced at the same time, synergistic cytotoxic effects are readily observed. A defense against nitrosative stress is provided by antioxidants and primarily by glutathione. Upon excess NO production, the antioxidant pool may be depleted, which will enhance nonspecific oxidation reactions. The toxic action has an important role, especially in nerve systems. It is assumed that during a stroke, for example, excess production of NO leads to the death of nerve cells. Direct modification of ligand-gated ion channels, such as the receptor for *N*-methyl-D-aspartate (NMDA receptor), by NO has a special role in connection with this.

Nitrosative stress, however, is also used in a beneficial way by the cell. It is increasingly recognized that nitrosative stress is an efficient tool by which the cell fights microbial infections. During infections, increased NO production is observed in humans and experimental animals, specifically in cells engaged in antimicrobial defense such

as phagocytes. Although the microbial targets of NO responsible for antimicrobial effects are poorly characterized, microbial proteins containing metal ion centers and reactive cysteine residues (e. g. cysteine proteases) are likely attack points.

Regulatory Function of NO

NO produced in a regulated way by enzymatic synthesis is involved in the control of a wide array of cellular functions including relaxation of blood vessels, neurotransmission, cellular immune response and apoptosis. Because of its high reactivity, NO can interact and react with many effector proteins.

Attack points are metal ion centers and specific cysteine residues of proteins. The mechanisms by which cysteine nitrosylation regulates protein functions can be broadly described in allosteric terms similar to protein phosphorylation. Often, O_2-mediated redox reactions cooperate in the allosteric control by NO of protein functions. S-nitrosylation of target proteins is a redox-based signal with exquisite specificity based on the selective modification of single cysteine residues. The selectivity of S-nitrosylation has been shown to be provided by both the subcellular distribution of NOS enzymes and the sequence context of cysteine residues in target proteins. Two nitrosylation motifs have been identified. In one motif, the target cysteine is located between an acidic and a basic amino acid, as revealed in either the primary sequence or the tertiary structure. In the other motif, the cysteine is contained in a hydrophobic region.

In Table 6.3, some important bioregulatory proteins are summarized, for which direct modification by NO has been shown. Two target proteins should be mentioned, in particular:

Tab. 6.3 Regulatory attack points of NO.
Proteins are included for which a direct regulation by NO is assumed (according to Stammler, 1994). Direct evidence of regulatory nitrosylation has only been shown for hemoglobin, however.

Binding site	**subcellular localization**			
	membrane	cytosol (incl. cpmpartments)	nucleus	extracellular
thiol	NMDA receptor NADPH oxidase protein kinase C adenylyl cyclase (type I)	aldolase GAPDH plasminogen activator aldehyde dehydrogenase	AP-1 NF_x B OMDM transferase	glutathione albumin
metal		guanylyl cyclase hemoglobin aconitase/IREP1 cyclooxygenase cyt P450		

Abbreviation: NMDA: N-methyl-D-aspartate; GAPDH: glycerine aldehyde-3-phosphate dehydrogenase; IREP1: iron responsive element binding protein; OMDM transferase: O^6-methylguanine-DNA methyltransferase

NO-sensitive Guanylyl Cyclase

The first cellular target of NO to be identified was a specific isoform of guanylyl cyclase (see Section 6.3). Stimulation of NO synthase leads to activation of a cytoplasmic NO-sensitive guanylyl cyclase. Activation is achieved by binding of NO to a heme group of the enzyme. The associated increase in the cGMP level has multiple consequences. The cGMP can stimulate cGMP-dependent protein kinases; it can also open cGMP-controlled ion channels. As a consequence, an increase in the intracellular Ca^{2+} concentration takes place and a Ca^{2+} signal is produced. NO can influence both protein phosphorylation and $InsP_3$/diacylglycerol and Ca^{2+} metabolism by this mechanism and activate a broad palette of biochemical reactions in the cell.

S-Nitrosylation of Hemoglobin

Hemoglobin was the first protein for which a regulatory action of S-nitrosylation was clearly shown (Lia et al., 1996; Fig. 6.19). Hemoglobin (Hb) is a tetramer, composed of

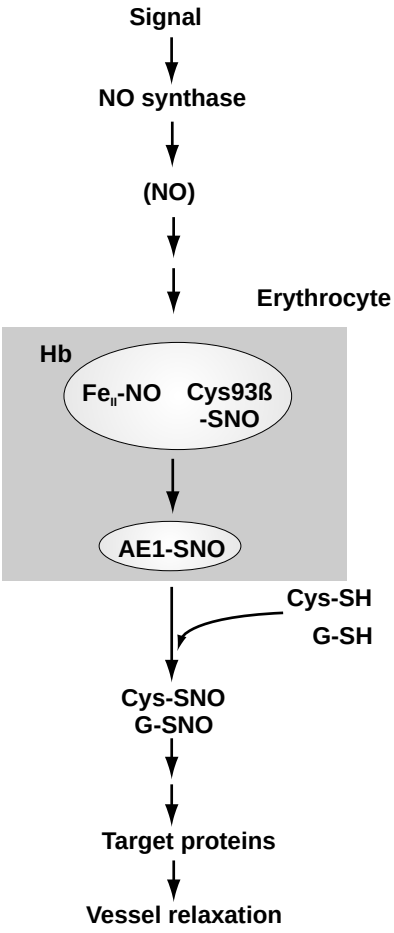


Fig. 6.19 Scheme of the function of nitroso-hemoglobin. NO synthase is activated by a stimulatory signal (e. g. a Ca^{2+} signal) and NO is formed. The NO is transferred by direct or indirect means to hemoglobin in the erythrocytes. NO can bind to hemoglobin as a Fe-NO complex with the heme, and it can exist as a S-nitroso derivative of Cys93 of the β subunit (Cys93β). Hemoglobin-bound NO can be transferred to the anion exchange protein AE1, forming SNO-AE1. This can transfer NO activity out of the red blood cell and into the vessel wall.
In this form, and transferred to low molecular weight SH compounds such as glutathione (GSH) or free cysteine (Cys). The resulting nitrosyl compounds Cys-SNO and G-SNO can diffuse to target proteins and pass the NO signal on to these. The figure does not show the complex regulation of NO compounds by hemeglobin via oxygen.

two α and two β chains. In man, each chain has a heme system, and the β chains have a reactive cysteine group (Cys93). The Hb may bind NO at two sites. Firstly, NO can bind to the Fe(II) of the heme grouping; secondly, NO can accumulate at Cys93 of the β chain by forming an S-nitrosyl.

Nitrosylation of hemoglobin is a dynamic vesicle function. The Hb functions as a means of transport which helps the erythrocytes to store NO and to transport it in the vessel system. The heme- and Cys93-bound NO can be delivered by Hb in a regulated way. The NO-binding capacity of hemoglobin is linked to O_2 binding to hemoglobin and is allosterically controlled. Binding of O_2 to hemoglobin and the associated transition from the T to the R form (see Section 2.3) facilitates binding as S-nitrosyl, whilst delivery of O_2 and the transition from the R to the T form leads to dissociation of NO. By this mechanism, selective delivery of NO to oxygen-depleted tissues is provided, resulting in blood vessel relaxation and increased blood flow.

How NO disembarks from the S-nitrosylated Hb and how it is delivered to the neighboring cells has long been not understood. There is now experimental evidence that this process requires a further protein that is present in abundant amounts in the cell membrane of red blood cells, the anion exchanger protein AE1 (review: Pawloski and Stamler, 2002). AE1 appears to be the primary recipient of NO groups from S-nitrosylated Hb and seems to accept the NO group directy from Hb, a step that is essential for the efflux of NO's biological activity from red blood cells.

Reference

1 Alderton, W.K., Cooper, C.E. and Knowles, R.G. (2001) Nitric oxide synthases: structure, function and inhibition. *Biochem.J.*, **357**, 593–615.

2 Ames, J.B., Ishima, R., Tanaka, T., Gordon, J.I., Stryer, L. and Ikura, M. (1997) Molecular mechanics of calcium-myristoyl switches. *Nature*, **389**, 198–202.

3 Anderson, R.A., Boronenkov, I.V., Doughman, S.D., Kunz, J. and Loijens, J.C. (1999) Phosphatidylinositol phosphate kinases, a multifaceted family of signaling enzymes. *J.Biol.Chem.*, **274**, 9907–9910.

4 Berridge, M.J. (1997) The AM and FM of calcium signalling. *Nature*, **386**, 759–760.

5 Bootman, M.D., Collins, T.J., Peppiatt, C.M., Prothero, L.S., Mackenzie, L., de Smet, P., Travers, M., Tovey, S.C., Seo, J.T., Berridge, M.J. *et al.* (2001) Calcium signalling–an overview. *Semin.Cell Dev.Biol.*, **12**, 3–10.

6 Bootman, M.D., Berridge, M.J. and Roderick, H.L. (2002) Activating calcium release through inositol 1,4,5-trisphosphate receptors without inositol 1,4,5-trisphosphate. *Proc.Natl.Acad.Sci.U.S.A*, **99**, 7320–7322.

7 Bootman, M.D., Berridge, M.J. and Roderick, H.L. (2002) Calcium signalling: more messengers, more channels, more complexity. *Curr.Biol.*, **12**, R563-R565.

8 Cantrell, D.A. (2001) Phosphoinositide 3-kinase signalling pathways. *J.Cell Sci.*, **114**, 1439–1445.

9 Chan, T.O., Rittenhouse, S.E. and Tsichlis, P.N. (1999) AKT/PKB and other D3 phosphoinositide-regulated kinases: kinase activation by phosphoinositide-dependent phosphorylation. *Annu.Rev.Biochem.*, **68**, 965–1014.

10 Chin, D. and Means, A.R. (2000) Calmodulin: a prototypical calcium sensor. *Trends Cell Biol.*, **10**, 322–328.

11 De Koninck, P. and Schulman, H. (1998) Sensitivity of CaM kinase II to the frequency of Ca2+ oscillations. *Science*, **279**, 227–230.

12 de Rooij, J., Zwartkruis, F.J., Verheijen, M.H., Cool, R.H., Nijman, S.M., Wittinghofer, A. and Bos, J.L. (1998) Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. *Nature*, **396**, 474–477.

13 Francis, S.H., Turko, I.V. and Corbin, J.D. (2001) Cyclic nucleotide phosphodiesterases: relating structure and function. *Prog.Nucleic Acid Res.Mol.Biol.*, **65**, 1–52.

14 Hausladen, A. and Stamler, J.S. (1999) Nitrosative stress. *Methods Enzymol.*, **300**, 389–395.

15 Hoeflich, K.P. and Ikura, M. (2002) Calmodulin in action: diversity in target recognition and activation mechanisms. *Cell*, **108**, 739–742.

16 Hofmann, F., Ammendola, A. and Schlossmann, J. (2000) Rising behind NO: cGMP-dependent protein kinases. *J.Cell Sci.*, **113 (Pt 10)**, 1671–1676.

17 Houslay, M.D. and Milligan, G. (1997) Tailoring cAMP-signalling responses through isoform multiplicity. *Trends Biochem.Sci.*, **22**, 217–224.

18 Irvine, R.F. and Schell, M.J. (2001) Back in the water: the return of the inositol phosphates. *Nat.Rev.Mol.Cell Biol.*, **2**, 327–338.

19 Irvine, R.F. and Schell, M.J. (2001) Back in the water: the return of the inositol phosphates. *Nat.Rev.Mol.Cell Biol.*, **2**, 327–338.

20 Ishii, M., Inanobe, A. and Kurachi, Y. (2002) PIP3 inhibition of RGS protein and its reversal by Ca2+/calmodulin mediate voltage-dependent control of the G protein cycle in a cardiac K+ channel. *Proc.Natl.Acad.Sci.U.S.A*, **99**, 4325–4330.

21 Kranenburg, O. and Moolenaar, W.H. (2001) Ras-MAP kinase signaling by lysophosphatidic acid and other G protein-coupled receptor agonists. *Oncogene*, **20**, 1540–1546.

22 Krugmann, S. and Welch, H. (1998) PI 3-kinase. *Curr.Biol.*, **8**, R828.

23 Pawloski, J.R. and Stamler, J.S. (2002) Nitric oxide in RBCs. *Transfusion*, **42**, 1603–1609.

24 Potter, L.R. and Hunter, T. (2001) Guanylyl cyclase-linked natriuretic peptide receptors: structure and regulation. *J.Biol.Chem.*, **276**, 6057–6060.

25 Ruvolo, P.P. (2001) Ceramide regulates cellular homeostasis via diverse stress signaling pathways. *Leukemia*, **15**, 1153–1160.

26 Stamler, J.S., Lamas, S. and Fang, F.C. (2001) Nitrosylation. the prototypic redox-based signaling mechanism. *Cell*, **106**, 675–683.

27 Testa, J.R. and Bellacosa, A. (2001) AKT plays a central role in tumorigenesis. *Proc.Natl.Acad.Sci.U.S.A*, **98**, 10983–10985.

28 Torreilles, J. (2001) Nitric oxide: one of the more conserved and widespread signaling molecules. *Front Biosci.*, **6**, D1161-D1172.

29 Vanhaesebroeck, B. and Alessi, D.R. (2000) The PI3K-PDK1 connection: more than just a road to PKB. *Biochem.J.*, **346 Pt 3**, 561–576.

30 Wedel, B. and Garbers, D. (2001) The guanylyl cyclase family at Y2K. *Annu.Rev.Physiol*, **63**, 215–233.

31 Wu, L., Niemeyer, B., Colley, N., Socolich, M. and Zuker, C.S. (1995) Regulation of PLC-mediated signalling in vivo by CDP-diacylglycerol synthase. *Nature*, **373**, 216–222.

32 Wu, X., Senechal, K., Neshat, M.S., Whang, Y.E. and Sawyers, C.L. (1998) The PTEN/MMAC1 tumor suppressor phosphatase functions as a negative regulator of the phosphoinositide 3-kinase/Akt pathway. *Proc.Natl.Acad.Sci.U.S.A*, **95**, 15587–15591.

7
Ser/Thr-specific Protein Kinases and Protein Phosphatases

Reversible phosphorylation of amino acid side chains is a widely used principle for regulation of the activity of enzymes and signaling proteins (see Chapter 3). Via this function, protein kinases and protein phosphatases play pivotal roles in regulating aspects of metabolism, gene expression, cell growth, cell division and cell differentiation. Almost all intracellular signaling pathways use protein phosphorylation to create signals and conduct them further. The protein kinases are certainly one of the largest protein families in the cell. Conservative estimates suggest that more than 1000 protein kinases are coded in the mammalian genome. Of the various protein kinases, the Ser/Thr-specific and Tyr-specific enzymes are the best characterized. Tyr-specific protein kinases are dealt with in Chapters 8 and 11. Before going on to the protein family of Ser/Thr-specific protein kinases, a rough classification of protein kinases will be presented.

7.1
Classification, Structure and Characteristics of Protein Kinases

7.1.1
General Classification and Function of Protein Kinases

The first protein kinase obtained in a purified form was the Ser/Thr-specific phosphorylase kinase of muscle, in 1959 (Krebs et al., 1959). With the discovery of the Tyr-specific protein kinases (Erikson et al., 1979), the Ser/Thr-specific protein kinases were joined by another extensive class of protein kinases of regulatory importance, to which a central function in growth and differentiation processes was soon attributed. At present, several hundred different protein kinases are known in mammals, most of which are Ser/Thr- or Tyr-specific. In addition, there are some protein kinases that phosphorylate other amino acids.

Based on the nature of the acceptor amino acids, four classes of protein kinases can be distinguished (Fig. 7.1):

– *Ser/Thr-specific protein kinases* esterify a phosphate residue with the alcohol group of Ser and Thr residues.

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

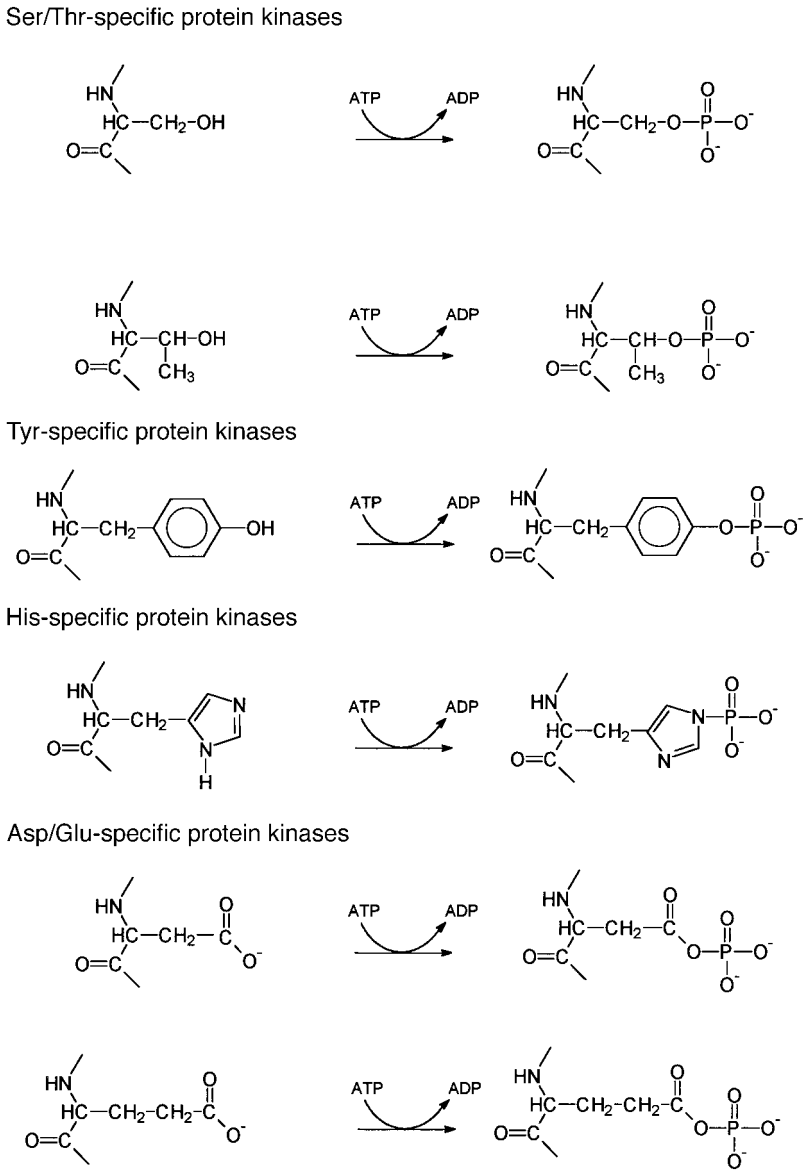


Fig. 7.1 Amino acid specificity of protein kinases.

– *Tyr-specific protein kinases* create a phosphate ester with the phenolic OH group of Tyr residues.
– *Histidine-specific protein kinases* form a phosphoro(?)us amide with the 1 or 3 position of His. The members of this enzyme family also phosphorylate Lys and Arg residues.

– *Aspartate- or glutamate-specific protein kinases* create a mixed phosphate-carboxylate anhydride.

Reversible phosphorylation of proteins on Ser/Thr and Tyr residues is a regulatory signal that functions as a switch in signaling pathways. The phosphate esters formed on proteins by the action of protein kinases are stable modifications that cause profound changes in the activity of cellular proteins. Because of the stability of the phosphate esters, protein phosphatases are required for their removal. The concerted and highly regulated action of both protein kinases and protein phosphatases is used by the cell to create a temporally and spatially restricted signal influencing the activity state of proteins in a highly specific way.

Examples of cellular activities regulated by protein kinases are diverse, affecting practically all the cell's performance. Protein phosphorylation is found in

– enzymes: as elements of signal chains (Chapters 7, 8, 10 and 13) and in enzymes of intermediary metabolism (Chapter 2)
– adaptor proteins (Chapter 8)
– signal proteins (Chapters 5, 7 and 8)
– transcription factors (Chapter 1)
– ion channels
– transmembrane receptors (Chapters 5, 8, 11 and 12)
– ribosomal proteins (ribosomal protein S6, Chapter 6)
– structural proteins
– transport proteins.

The switch function of protein phosphorylation is based on different mechanisms that may work alone or in combination. Phosphorylation by protein kinases influences function and activity of the protein substrate, especially in the following ways:

– Induction of conformational changes by allosteric mechanisms (Chapters 2, 11 and 13)
– Direct interference with binding of substrate or other binding partners, e. g., isocitrate dehydrogenase (Chapter 2)
– Creation of binding sites for effector molecules in the sequence: examples of this are binding of Tyr-P to SH2 and PTB domains and binding of Ser-P to 14-3-3 proteins (Chapter 8).

By far the most data are available on the structure and function of the Ser/Thr- and Tyr-specific protein kinases, and the following discussion will therefore concentrate on these two protein kinase classes.

In the sequence of protein kinases, a homologous catalytic domain can be identified that includes ca. 270 amino acids. Based on the sequence of the catalytic domain, it is possible to differentiate between Ser/Thr- and Tyr-specific protein kinases. Furthermore, homology considerations enable subfamilies within both the larger families of Ser/Thr- and Tyr-specific protein kinases to be identified.

7.1.2 Classification of Ser/Thr-specific Protein Kinases

Because of the complexity and number of the subfamilies, only selected subfamilies of the Ser/Thr-specific protein kinases in vertebrates are presented in the following:

Subfamilies of the Ser/Thr-specific protein kinases

- Protein kinases regulated by a cyclic nucleotide
  - cAMP regulated protein kinase, protein kinase A, PKA
  - cGMP regulated protein kinase
- Diacylglycerol-regulated protein kinases
  Protein kinase C, with further division into the α, βI, βII, γ, δ, ε, ζ, η, , ι, λ, and μ subtypes
- Calcium/calmodulin-regulated protein kinases
  - subunit of phosphorylase kinase
  - Myosin light chain kinase, MLCK
  - Ca^{2+}/calmodulin dependent protein kinase II
- Ribosomal S6 protein kinase.
  Kinases that specifically phosphorylate ribosomal protein S6
- G protein-coupled receptor kinase
- β-adrenergic receptor kinase, βARK
- Casein kinase II (casein kinase gets its name from the observation that casein, milk protein, is a good substrate)
- Glycogen synthase kinase
- CDC2 kinases (representatives of this family are central elements of regulation of the cell cycle, see Chapter 143)
- Mitogen-activated kinases, MAP kinases (the MAP kinases are involved in the transduction of growth-promoting signals, see Chapter 10)
- Mos/Raf protein kinases (see Chapter 9; the Mos/Raf protein kinases are also involved in signal transduction of growth factors).

There are many other protein kinases that do not show any close relationship to these subfamilies. These include protein kinases with twofold specificity in that they can phosphorylate Ser/Thr and also Tyr residues. An example of a protein kinase with twofold specificity is the MAP kinase kinase (see Chapter 10).

We also know of Ser/Thr-specific protein kinases that are an integral part of transmembrane receptors or of ion channels. The TGFβ-receptor contains a Ser/Thr-specific protein kinase activity in the cytoplasmic part of its transmembrane polypeptide chain (see Chapter 12). Some members of a certain class of ion channels, named the transient receptor potential channels (TRP channels), carry a protein kinase activity on the cytoplasmic side of the channel protein that is essential for channel function. The TRP channels modulate Ca^{2+} entry into eukaryotic cells in response to external signals. The protein kinase activity is located on the cytoplasmic domain of the channel and can phosphorylate itself and other proteins on Ser/Thr residues (Runnels et al., 2001).

Although there is no similarity to classical protein kinases on the primary sequence level, the three-dimensional structure of the TRP channel protein kinase domain is very similar to the classical kinase fold.

7.2
Structure and Regulation of Protein Kinases

The Ser/Thr- and Tyr-specific protein kinases share many common features. The catalytic mechanism, structural properties and mechanisms of kinase control are very similar for the two kinase classes. In the following, the main properties of both classes will therefore be presented together (reviews: Johnson et al., 1998; Engh and Bossemeyer, 2001) .

The Protein Kinase Reaction

The common catalytic function of protein kinases is the covalent phosphorylation of substrate proteins via transfer of the γ-phosphate of ATP to the OH group of serine, threonine or tyrosine residues. This catalytic function is carried out by a catalytic do-

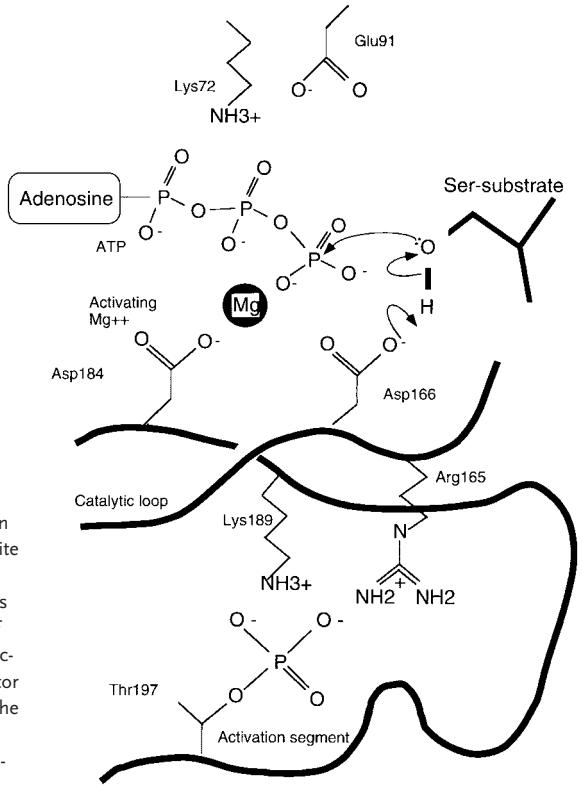


Fig. 7.2 Schematic representation of key interactions at the catalytic site and the activation loop of protein kinase A. A possible mechanism is shown in which the carboxylate of Asp166 functions as a base and activates the OH-group of the acceptor serine for a nucleophilic attack on the γ-phosphate of ATP: The catalytic mechanism is not definitively established.

main of ca 270 amino acids whose structure and catalytic residues are highly conserved among the protein kinase family. The conserved structure shows two lobes that are linked by a flexible hinge. The catalytic center is formed by residues from both lobes. The proposed mechanism of phosphate transfer by protein kinases is shown in Fig. 7.2 (see Johnson et al., 1996). It should be kept in mind, however, that this mechanism is not definitely established.

Sequence comparisons, mutation experiments and biochemical studies indicate an essential function in the catalysis of phosphate transfer for the conserved amino acids Lys72, Asp166 and Asp184 (numbering of PKA). It is generally assumed that Asp166, which is invariant in all protein kinases, serves as a catalytic base for activation of the Ser/Thr hydroxyl and that the reaction takes place by an "in-line" attack of the Ser-OH at the γ-phosphate.

Key players of the reaction are

- acidic amino acids required for
  - stabilization of the transition state
  - activation of the OH group of the acceptor amino acid for nucleophilic attack on the γ-phosphate
- one or two metal ions that coordinate the γ-phosphate of ATP, help to fix the ATP, and stabilize negative charges in the transition state
- basic amino acids that serve to stabilize negative charges in the transition state and in the leaving group ADP.

Overall, the mechanism of phosphate transfer by protein kinases is related to nucleotide transfer by nucleic acid-polymerizing enzymes that also use metal ions and acidic residues as key elements in catalysis.

7.2.1 Main Structural Elements of Protein Kinases

The core of all eukaryotic Ser/Thr- and Tyr-specific protein kinases adopt a common fold illustrated in Fig. 7.3 for the tyrosine kinase domain of the insulin receptor. The structure comprises two lobes that are connected by a hinge region. The N-terminal lobe contains five β-structures and one α-helix, named C-helix. In contrast, the larger C-terminal lobe is mostly α-helical. It comprises a four-helix bundle, additional α-helices, and two short β-strands. ATP and 1 or 2 metal ions are bound at the interface of the two lobes, while the binding site for the peptide substrate is located mostly in the C-terminal lobe. The following structural elements have been found to be critical for catalysis and for protein kinase control:

- **Glycine-rich loop in the N-terminal lobe**
  This loop (also called P-loop) is found in similar form also in the G_α-proteins (see Section 5.4.3) and is required for anchoring of the phosphate residues of ATP.
- **C-helix of the N-terminal lobe**
  In most active protein kinase conformations, the C-helix forms a salt bridge to an invariant Lys-residue within the N-terminal lobe, allowing optimal positioning of

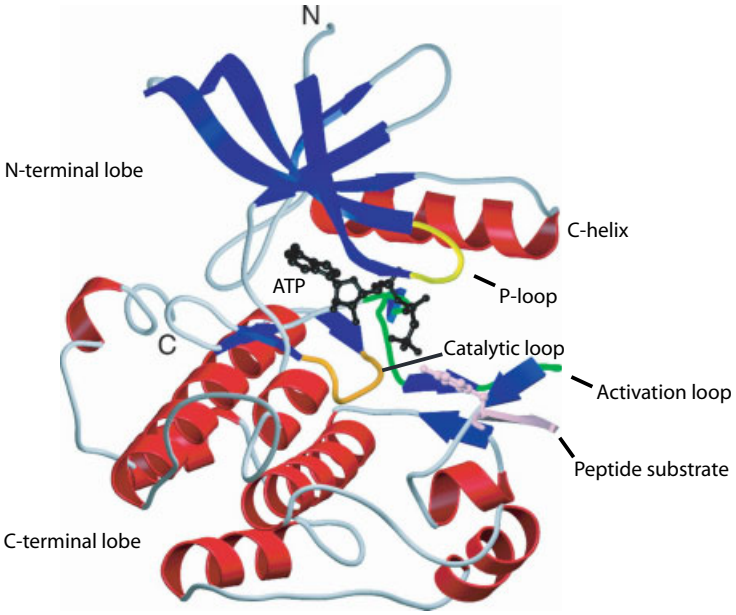


Fig. 7.3 Structure of the tyrosine kinase domain of the Insulin receptor with bound ATP and peptide substrate (after Hubbard and Till, 2000). The α-helices are shown in red, the β-strands in blue. The functional important structural elements are indicated. The P-loop is shown in yellow, the activation loop in green and the catalytic loop in orange. The amino and carboxy termini are denoted by N and C.

the ATP phosphates. In many inactive conformations the conformation of the C-helix is changed and the salt bridge is absent.

- **Catalytic loop**
  A conserved Asp-residue (Asp 166 of protein kinase A), presumed to be the catalytic residue, and a conserved Asp-residue involved in Mg^{2+}-ATP binding, are found on this loop.
- **Activation segment**
  The activation segment (also called T-loop for the cyclin-dependent protein kinases, see Chapter 13) comprises 19 – 32 amino acids and contains a conserved DFG motif (single letter code). The D residue is involved in metal-ATP binding. In many, but not all protein kinases, Ser, Thr or Tyr residues of the activation loop are phosphorylated in response to activating signals. This phosphorylation promotes an active conformation of the kinases. By this token, the activation segment is one of the most important control elements of protein kinase activity. Phosphorylation of the activation segment serves to create the protein substrate-binding site and/or to reorient the catalytic residues for optimal phosphate transfer.
- **Autoinhibitory sequence elements**
  Many protein kinases contain autoinhibitory elements that help to fix an inactive conformation by intramolecular binding to the substrate-binding site. These ele-

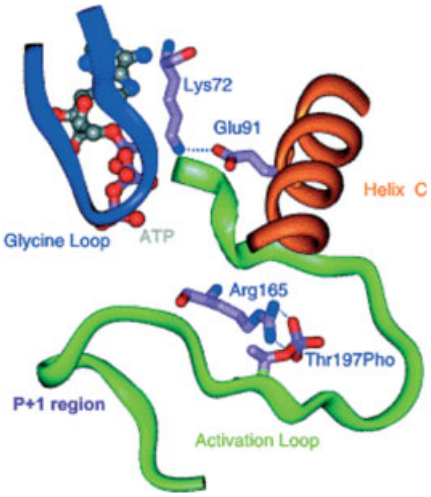


Fig. 7.4 Alignment of the activity modulation sites of protein kinases (after Hubbard and Till, 2000). Close-up of the critical elements of protein kinase control as found in active kinase structures, represented here by the structure of active protein kinase A. The phosphothreonine acts as an organizing center that forms contacts to Arg 165 which is neighboured to the katalytic D166 and to the C-helix. The C-helix forms a salt bridge to K72 which helps to coordinate the α- and β- phosphates of ATP correctly. This helix corresponds to the PSTAIRE helix of CDK2 (see Chapter 13) and is the focus of several regulatory mechanisms among protein kinases. In inactive kinase structures, two or more of the critical activity modulation sites are misaligned or blocked by intra- or extramolecular interaction.

ments occupy the substrate peptide-binding groove in a mode like the real substrate; however, they lack the phosphorylatable residues and are therefore also termed pseudo-substrates.

In inactive kinase structures, two or more of the above mentioned structural elements are aligned in a way that prevents substrate binding and/or catalysis. Fig 7.4 shows the spatial relationship of the critical protein kinase activity modulation sites.

7.2.2 Substrate Binding and Recognition

Taking into account the many Ser, Thr and Tyr residues in proteins, the question arises of which parameters define the phosphorylation site of a substrate protein. With the help of targeted exchange of amino acids in substrate proteins, sequence comparison of phosphorylation sites, and use of defined peptides as substrates, it has been possible to clearly show that the sequence in the neighborhood of a Ser/Thr or Tyr residue is an important determinant of specificity. The different Ser/Thr- and Tyr-specific protein kinases show different requirements with respect to the neighboring sequence of the residues to be phosphorylated, so that each subfamily has its own consensus sequence for phosphorylation. Within the family of Ser/Thr kinases, the sequences of the phosphorylation sites show great variability, and it is not possible to give a consensus sequence for all Ser/Thr-specific protein kinases. Because of the different roles and the very different substrate proteins, this is not surprising. It should be pointed out that several Ser/Thr residues are found in many phosphorylation sequences, so that multiple and cooperative phosphorylation is possible in a sequence segment. Phosphorylation of the large subunit of the RNA polymerase II (see Section 1.4.3) is particularly marked. At the C terminus,

this contains 52 copies of the heptamer sequence YSPTSPS as potential phosphorylation sites.

The consensus sequence is, however, not the only determinant of substrate specificity. When assayed *in vitro*, many protein kinases show indistinct substrate specificity, whereas *in vivo* only certain substrates are phosphorylated, indicating that additional factors are relevant to substrate selection *in vivo*. Structural parts of the substrate protein that are distinct from the phosphorylation site have been shown to contribute to substrate binding as well. A coupling of a protein kinase and its protein substrate can be achieved, e. g., via structural parts that are located far away from the phosphorylation consensus site on the substrate and the substrate binding site on the kinase (review: Biondi and Nebrada, 2003). An example is provided by the docking of substrate proteins on receptor tyrosine kinases via SH2-phosphotyrosine interactions (see Chapters 8 and 11). This interaction helps to clamp the substrate on the kinase, thereby ensuring a high efficiency of phosphate transfer. Another major determinant of protein kinase specificity is the targeting of protein kinases to the neigborhood of selected substrates. The colocalization of protein kinases and their substrates at distinct subcellular sites greatly enhances the specificity of the kinase reaction. Only those substrates that have been translocated to the specific subcellular site will be phosphorylated by the protein kinase. The mechanisms of colocalization are diverse and will be dealt with separately in Section 7.7.

The peptide substrate-binding site is located on the C-terminal lobe of the protein kinases. Structural information available to date shows that the peptide substrate is contacted via multiple interactions, both N- and C-terminal, to the residue to be phosphorylated. There is a marked complementarity between the binding pocket on the kinase and the peptide substrate in regard to shape, hydropathy and electrostatic potential. For several protein kinases as, e. g., the insulin receptor tyrosine kinase (see Section 8.1.3) and CDK2 (see Section 13.2.5), an intact peptide-binding site is not present in inactive forms of the kinase. Only after activation segment phosphorylation and a subsequent conformational change is a substrate recognition site created in these kinases.

7.2.3
Control of Protein Kinase Activity

Protein kinases can exist in active and inactive forms, which is why they are able to perform the function of a switch in signaling pathways. For most of the time, protein kinases are found in the inactive, "off" form. Upon specific signals they are converted into the active "on" form and can now phosphorylate substrate proteins. The duration and extent of activation depends on the nature of the signals that induce the transition from the inhibited state to the activated state. Termination of the activated state may be achieved, e. g., via dephosphorylation of the activation loop by protein phosphatases.

The cell uses a multitude of mechanisms for control of the activity of protein kinases. These controls operate mainly at two levels :

– via allosteric regulation of kinase activity
– via colocalization of kinase and substrate protein.

Overall, the protein kinase structures contain several flexible elements that can be fixed in either an active or an inactive conformation. The flexible hinge between the two lobes allows for their regulated movement. Other highly mobile elements are the activation segment and, to a lesser extent, the C-helix. The structural information on the active state of several protein kinases shows a very similar structure of the catalytic domain. The following structural features are characteristic of the activated state of protein kinases:

– The two lobes are closely packed together
– The ATP is buried between the two lobes
– The C-helix is positioned for salt bridge formation to a β-strand of the N-terminal lobe
– The substrate binding site is fully formed
– The phosphorylated activation segment helps to organize the catalytic residues for optimal phosphate transfer.

Whereas fully active protein kinases appear to adopt a similar active conformation, the inactive states of protein kinases are very diverse. There are many ways to fix the flexible protein kinase structure in an inactive state. Overall, the inactive states are characterized by a more open conformation of the two lobes, precluding optimal orientation of residues involved in substrate binding and catalysis. The cell uses the following mechanisms for fixation of inactive protein kinase states (Fig.7.5):

- Binding of protein inhibitors: protein kinase inhibitors fix inactive states by, e. g., deforming the N-terminal lobe and destroying the ATP-binding site; see inhibitor $p21^{Kip}$, Section 13.2.6
- Inhibitory phosphorylations (e. g., Thr14, Tyr15 on CDKs, see Chapter 13, phosphorylation of Src-kinase, see Section 8.3.2)
- Binding of regulatory subunits (see protein kinase A, Section 7.3)
- Autoinhibition: if an inhibitory structural element is itself part of the protein kinase, this is known as *autoinhibition*. The inhibitory structural elements often have the character of a pseudosubstrate. They possess a similarity to the proper substrate and can accumulate in the substrate-binding site. Since the pseudosubstrate lacks a phosphate receptor, no phosphorylation takes place. Autoinhibition of twitchin kinase, which is a Ser/Thr-specific protein kinase of the nematode *Caenorhabditis elegans* and is homologous to the myosin light-chain kinase of mammals, is based on several structural alterations (Hu et al., 1994). The active site of twitchin kinase is blocked by the autoinhibitory structural element by
  – forming contacts with the substrate-binding site (pseudosubstrate function)
  – resulting interference with ATP binding
  – contacting essential catalytic residues and shielding these.

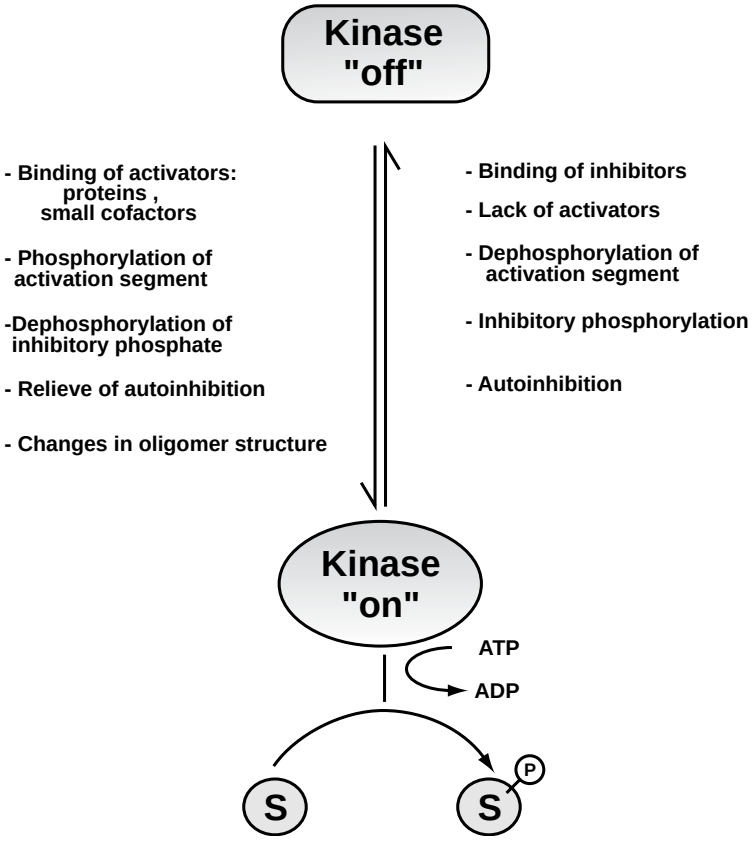


Fig. 7.5 Mechanisms of activation and inactivation of protein kinases.

A multitude of mechanisms are available to induce the transition from the inactive to the active state:

- Binding of activating subunits (cyclins, see Section 13.2.3)
- Binding of chemical messengers (cAMP, see Section 7.3.1) with concomitant release of inhibitory subunits
- Binding of cofactors like Ca^{2+}/ calmodulin, diacylglycerol, phospholipids (see Section 7.4)
- Phosphorylation of the activation loop
- Dephosphorylation of inhibitory phosphorylated sites
- Changes in the oligomerization state of the kinase as a consequence of ligand binding to the extracellular domain of the transmembrane protein kinase (see Section 8.1.2).

Generally, these modifications have the character of activating signals that are part of diverse signaling pathways and are used to recruit protein kinases as signaling elements.

7.3 Protein Kinase A

The cAMP-dependent protein kinase (protein kinase A, PKA) is the principal target of the second messenger cAMP. Of the protein kinases, protein kinase A is the best investigated and characterized (review: Skalhegg and Tasken, 2000). The substrates and functions of protein kinase A are diverse. Protein kinase A is involved in the regulation of metabolism of glycogen, lipids and sugars. In addition, cAMP/protein kinase A plays an important role in controlling ion channels and in long-term modifications at nerve synapses. Furthermore, it is involved in cAMP-stimulated transcription of genes that have a cAMP-responsive element in their control region (review: Mayr and Montminy, 2001). An increase in cAMP concentration leads to activation of protein kinase A, which phosphorylates the transcription factor CREB at Ser 133. CREB only binds to the transcriptional coactivator CBP in the phosphorylated state and stimulates transcription of target genes (see Section 1.4.5.2).

7.3.1 Structure and Substrate Specificity of Protein Kinase A

The activity of protein kinase A is primarily controlled by cAMP. In the absence of cAMP, protein kinase A exists as a tetramer composed of two regulatory R subunits and two catalytic C subunits (see Fig. 6.2). The catalytic activity is masked in the holoenzyme C_2R_2, since an inhibitory structural element of the R subunit blocks the entrance to the active site. Binding of cAMP to the R subunit leads to a reduction in the affinity between R and C by a factor of 10,000 – 100,000. The holoenzyme dissociates into the dimer of the R subunits and two monomers of C, which now become catalytically active.

In mammals, four isoforms of the R subunit (RIα, RIβ, RIIα and RIIβ) and three subtypes of the C subunit, namely Cα, Cβ and Cγ, are known. The existence of multiple R and C subunits harboring different biochemical features allows for the formation of a number of holoenzymes with different biological characteristics, which certainly contributes to the specificity and variablity of protein kinase A signaling observed in the cell.

The composition of the subunits is shown schematically in Fig. 7.6. The R subunit has two cAMP binding sites of differing affinity. In addition, the RII subunit has a domain containing an autophosphorylation site which is involved in the autoinhibition of protein kinase A.

The C subunit has a myristinic acid residue at the amino terminus, the function of which is unknown. In addition, the C subunit has specific Ser/Thr phosphorylation

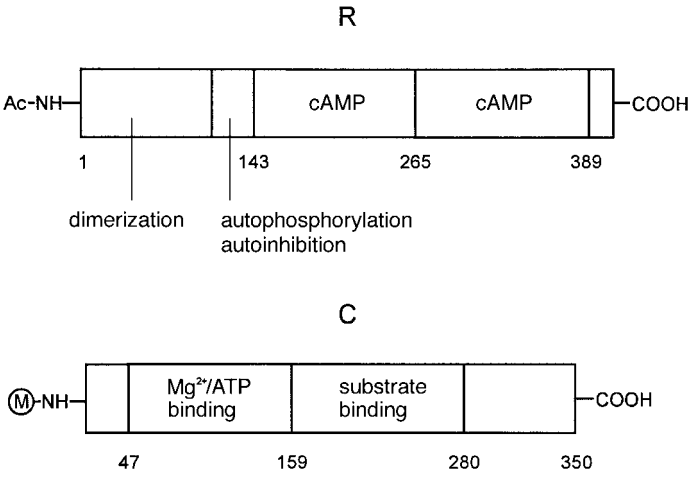


Fig. 7.6 Functional domains of protein kinase A. The functional domains of the catalytic (C) and regulatory (R) subunits of protein kinase A (bovine) are shown in a linear configuration. M: myristoylation.

sites, namely Thr197 and Ser338. Thr197 is located in the activation segment and is phosphorylated by an autophosphorylation mechanism. This is linked to an increase in the affinity for ATP and in catalytic efficiency.

The *consensus sequence* for phosphorylation of proteins by protein kinase A is RRXSX. The RII subunit contains such a sequence in the autoinhibitory domain and is therefore subject to phosphorylation by the C subunit in the holoenzyme, but without release of inhibition. Inhibition of the C subunit by the R subunit is based on binding of the autoinhibitory sequence of R at the substrate-binding site and at parts of the active center of the C subunit.

7.3.2 Regulation of Protein Kinase A

Protein kinase A is subject to multiple regulatory influences. Whereas the regulation by cAMP is the primary determinant of protein kinase A regulation, other regulatory influences ensure a high specificity of protein kinase A in regard to tissue distribution and compartment-specific action. Furthermore, the activity of protein kinase A can be controlled by the following mechanisms, both in a highly temporal and a spatially regulated way in the cell.

- **Changes in cAMP concentration**
  The changes in concentration of cAMP that lead to activation of protein kinase A in the cell are relatively small. In many tissues, a two- to threefold increase in cAMP concentration is sufficient to bring about the maximum physiological effect. The cell has different mechanisms available that limit the increase in cAMP concentra-

tion to a relatively narrow concentration region and contribute to damping of signal transduction via protein kinase A. An example of a mechanism with a damping effect in signal transduction by protein kinase A is a *feedback control* by a cAMP phosphodiesterase. The activated protein kinase A phosphorylates and activates a phosphodiesterase that hydrolyzes cAMP to AMP (Fig. 7.7). This mechanism enables protein kinase A to control its own steady-state activity. It also ensures that, as the external signal diminishes, the cAMP signal rapidly subsides.

- **Phosphorylation on Ser/Thr residues**
  Phosphorylation is another mechanism of protein kinase A control. Thus, the C subunit has been found to be specifically phosphorylated at two sites, whereby only one of the two phosphorylation events appears to be due to autophosphorylation. It is not clear, however, which signals induce the phosphorylation of the C subunit.

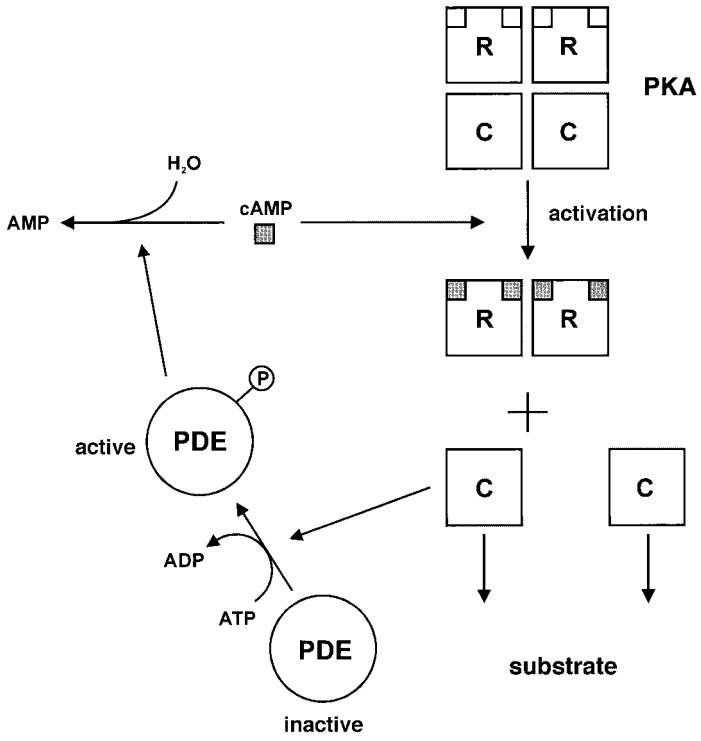


Fig. 7.7 Feedback control of protein kinase A by a phosphodiesterase. On activation of protein kinase A, the catalytic C subunits are released, which then phosphorylate a phosphodiesterase, in addition to other substrates. The phosphodiesterase is activated by the phosphorylation and hydrolyzes cAMP to AMP, whereby the signal transduction via protein kinase A is reduced or terminated.

- **Binding of inhibitor proteins**
  Regulation of protein kinase A may also take place via the binding of specific inhibitor proteins. The inhibitor PKI, for example, is involved in subcellular transport of C subunits and is considered a major regulator of C subunit activity.
- **Targeting to subcellular sites by A-kinase anchor proteins (AKAPs)**
  Given the fact that protein kinase A is widely distributed and is activated by a large variety of external signals, the question arises how – apart from isoenzyme patterns, differential phosphorylation and inhibitor binding – the high specificity of protein kinase A action is achieved in the cell. Specific compartmentalization of protein kinase A enzymes has now been recognized to be a major determinant of protein kinase A specificity. By binding to subcellular structures via A kinase-anchoring proteins (AKAPs), isoenzymes of proteinkinase A can be assembled at distinct subcellular sites in the vicinity of their substrates (review: Feliciello et al., 2001). We know of >20 different AKAPs that are located to different compartments of the cell (see Section 7.7). Most of the interactions between AKAPs and protein kinase A holoenzymes seem to be mediated by RII subunits.

7.4 Protein Kinase C

7.4.1 Characterization and Classification

The family of protein kinase C enzymes includes Ser/Thr-specific protein kinases, which were first identified by the requirement of the following cofactors (review articles: Oancea and Meyer, 1998; Mellor andParker, 1998. ; Dempsey et al., 2000):
– diacylglycerol
– Ca^{2+}
– phospholipids such as phosphatidylcholine.

Furthermore, protein kinase C enzymes have been shown to be high-affinity receptors for phorbol esters (see below). Regulation by Ca^{2+} and diacylglycerol identify protein kinase C enzymes as components of signal transduction pathways, in the course of which, phospholipase C is activated and the messenger substances $Ins(3,4,5)P_3/Ca^{2+}$ and diacylglycerol are produced. Therefore, activation of protein kinase C may take place via two central pathways:

Signaling pathways starting from receptor tyrosine kinases trigger stimulation of protein kinase C by activating phospholipase Cγ. An activating signal may also be despatched in the direction of protein kinase C – via activation of phospholipase Cβ – from G-protein-coupled membrane receptors (see Fig. 6.5f).

The sensitivity to Ca^{2+} and diacylglycerol and specific binding of phorbol esters has been considered for a long time to be the main characteristics of protein kinase C enzymes. Molecular cloning techniques and biochemical work, however, revealed that protein kinase C is a large kinase family that includes a variety of isoenzymes with very different regulatory properties (see below).

Stimulation by Phorbol Esters

A property of the protein kinase C enzyme family that is highly valuable for their identification and characterization is their activation by tumor promoters such as *phorbol esters* (Fig. 7.8). Protein kinase C binds to the tumor promoter, tetradecanoyl phorbol acetate (TPA), with high affinity and is activated by this binding. The specific activation of protein kinase C by phorbol esters is an important tool for demonstrating their involvement in signal transduction pathways. By external addition of TPA, it is possible to use model systems to test which biological responses of a signal transduction pathway involve, and are controlled by, protein kinase C.

Tumor promoters such as TPA do not themselves initiate tumor formation, but rather they promote triggering of the tumor by carcinogenic substances, e. g., benzo[a]pyrene. At present, the mechanism by which the structurally very heterogeneous tumor promoters exert their tumor-promoting activity is not understood. In the case of TPA, stimulation of protein kinase C forms the basis of the explanation. Since one of the roles of protein kinase C is in regulation of proliferation and of differentiation processes, unregulated activation of protein kinase C could lead to undesired protein phosphorylation and thus bring about misregulation of cell proliferation.

In addition to protein kinase C enzymes, cells contain other receptors for phorbol esters and diacylglycerol. In mammals, proteins named α- and β-chimaerins have been identified that lack protein kinase activity and bind phorbol esters and diacylglycerol with high affinity. Therefore, not all biological responses elicited by phorbol ester treatment can be attributed to protein kinase C stimulation (review: Ron and Kazanietz, 1999).

The Protein Kinase C Family

Like most of the Ser/Thr-specific protein kinase family, the protein kinase C family shows significant heterogeneity. At the present time, at least 12 different subtypes of protein kinase C, based on different criteria such as sequence, stimulation and regulation (Fig.7.9), have been discovered in mammals. These subtypes are grouped into

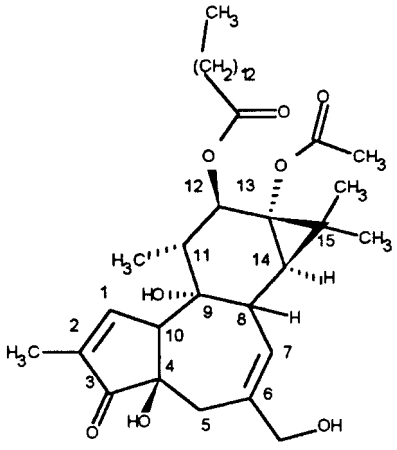


Fig. 7.8 Structure of tetradecanoyl phorbol acetate (TPA). Tetradecanoyl phorbol acetate functions as a tumor promoter and is a specific activator of protein kinase C.

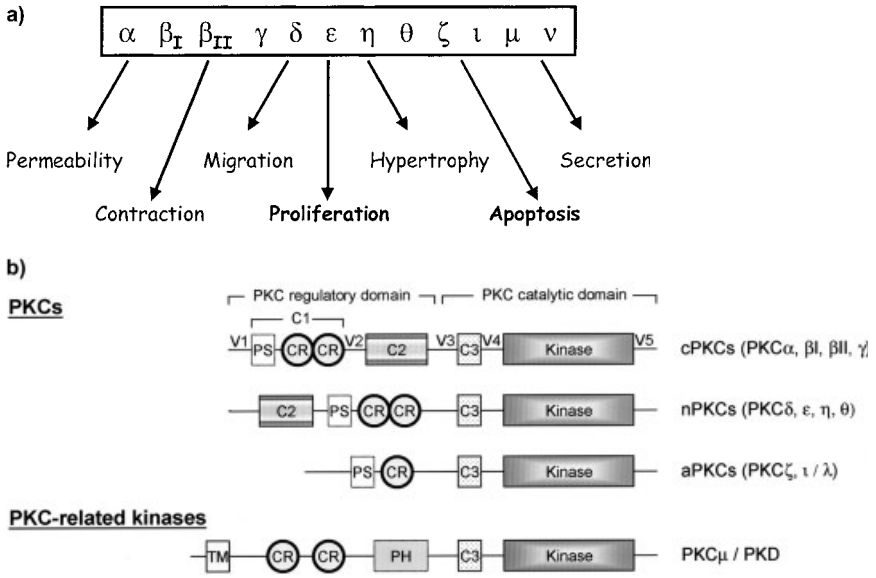


Fig. 7.9 a) Protein kinase C isoenzymes. The major cellular functions of the protein kinase C isoenzymes are indicated (from Dempsey et al., 2000). The isoenzymes are divided into subgroups based on structure and cofactor requirements: conventional (α, βI, βII,γ) novel (δ, ε, η, θ) atypical (ξ,ι) and protein kinase C related (μ, ν). b) Domain structure of protein kinase C isoenzymes. The phorbol esters and the second messenger diacylglycerol bind to the cysteine-rich motif present in conventional PKCs (cPKC) and novel PKCs (nPKC). The nPKCs are Ca^{2+} independent. V: variable region; PS: pseudosubstrate region; CR: cysteine-rich domain; TM: transmembrane domain; PH: pleckstrin homology domain.

the classical or conventional PKCs (cPKCs, subtypes α, βI, βII, γ), novel PKCs (nPKCs, subtypes δ, ε, θ, η), atypical PKCs (subtypes ζ, , ι), and the recently described ν and μ subtypes. Each isoenzyme is the product of a separate gene, with the exception of βI, βII enzymes, which are alternatively spliced variants of the same gene.

The subtypes have very different cofactor requirements. The "classical" PKCs are activated by Ca^{2+} and/or diacylglycerol and phorbol esters, while the "novel" PKCs are Ca^{2+}-independent and can be activated by diacylglycerol and phorbol esters. In contrast, the atypical PKCs are unresponsive to Ca^{2+} and diacylglycerol.

In addition to the different cofactor requirements, the PKC subtypes are distinguished by different cellular localization and a different pattern of substrate proteins. For example, the α, δ and ζ subtypes are widespread in almost all tissues, whereas the other subtypes only occur in specialized tissues.

The members of the protein kinase C family are composed of a polypeptide chain with a molecular weight of 68–83 kDa. The N-terminal regulatory domains C1 and C2 and a C-terminal catalytic domain can be differentiated in the primary structure (Fig. 7.9b) of the conventional PKCs. In addition, a pseudosubstrate sequence with autoinhibitory function is located in the C1 region that binds to the substrate-binding

site of the catalytic domain and keeps the enzyme in an inactive state in the absence of cofactors and activators.

The C1 domain is present in all PKC isoenzymes. It contains two cysteine-rich motifs (also known as Cys1- and Cys2 elements), each with two bound Zn^{2+} ions, which are involved in the binding of diacylglycerol and phorbol esters. When expressed in the isolated form, both cysteine-rich domains can specifically bind phorbol esters, whereas *in vivo* only one of the two binding sites is occupied by phorbol esters.

In the C2 section, which is not present in all members of the protein kinase C family, the Ca^{2+}-binding site and a binding site for acidic phospholipids are found. The conserved C3 and C4 sections form the catalytic domain with the binding sites for ATP and substrate proteins.

Most of the biochemical information is available for protein kinase C*α*. In the following, protein kinase C refers to the protein kinase C*α* enzyme.

7.4.2 Structure and Activation of Protein Kinase C

In the absence of activating cofactors, the catalytic domain is subject to *autoinhibition* by the regulatory domain. A sequence motif is found in the regulatory domain, which serves as a pseudosubstrate. It resembles the consensus sequence for phosphorylation sites of protein kinase C, but does not have a Ser or Thr residue for phosphorylation. This sequence motif is found in all protein kinase C family members. It is assumed that the active center is inhibited by occupation by the pseudosubstrate.

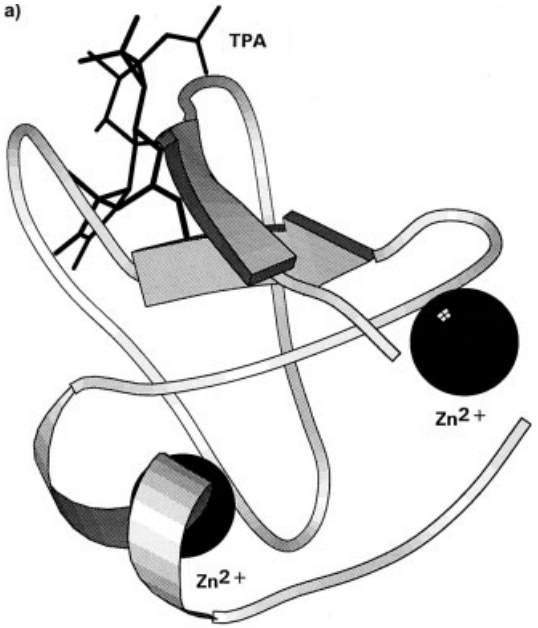
Two functions are attributed to the binding of the activating cofactors Ca^{2+}, diacylglycerol and phospholipid:

- stabilization of a structure of protein kinase C in which the active center is accessible for substrate proteins
- promotion of membrane association.

Detailed structural information on protein kinase C is not available at present, probably because of its high flexibility. The first insight into the mechanism of activation was obtained by structural determination of the cysteine-rich element (Cys2 element) of protein kinase C*δ* in complex with phorbol ester (Zhang at al., 1995).

From the structure of the Cys2 element with bound phorbol ester (Fig. 7.10a), it was concluded that the activating function of the phorbol ester is based, in particular, on promotion of membrane association of protein kinase C. The binding site of the phorbol ester lies in a hydrophobic region of the Cys2 element which is broken by a hydrophilic region. On binding of the phorbol ester, a continuous hydrophobic surface is created in this region of Cys2. It is assumed that the Cys2 section is involved in membrane association and that membrane association is energetically unfavorable in the absence of the phorbol ester. According to this theory, the phorbol ester binding enlarges the hydrophobic surface of the Cys2 element and favors a partial insertion of the Cys2 in the membrane (Fig. 7.10b).

Competition experiments have clearly shown that the second messenger diacylglycerol binds to the phorbol ester-binding site on the cysteine-rich domains in the C1



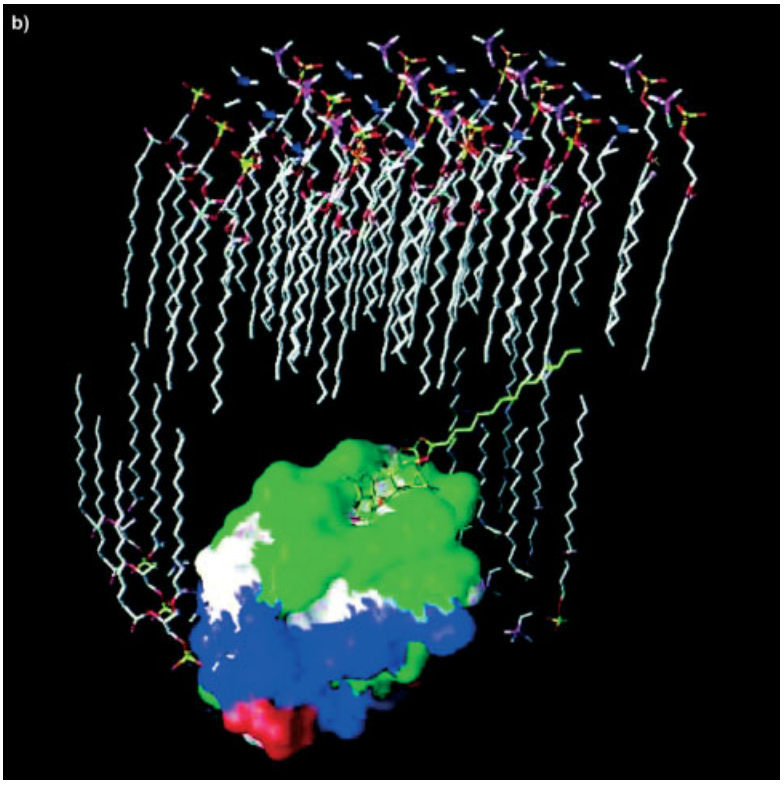


Fig. 7.10 a) Structure of the Cys2 element of protein kinase Cδ with bound phorbol ester. (Zang et al., 1995).
b) Model for the second cysteine-rich domain of protein kinase C complexed with phorbol ester (after Ron and Kanietz, 1999). Positively charged residues are shown in blue, negatively charged residueses are red, hydrophobic residues are green and neutral polar residues are white. The phorbol ester binds at the tip of the domain forming a hydrophobic cap over polar groups. By forming this continuous hydrophobic surface, the phorbol ester helps to insert PKC into the lipid bilayer.

region. It is generally assumed that this binding promotes membrane association of the PKC.

The C2 domain has been identified as the binding site for Ca^{2+}, phospholipids and proteins. Ca^{2+} and phospholipid binding to the C2 domain enhances membrane association and helps to release the enzyme from the autoinhibited state. Although first described for the PKC enzymes, the C2 domain has been recognized as a widespread domain also found in other enzymes like phosholipase C isoenzymes and phospholipase A_2.

7.4.3 Regulation of Protein Kinase C

The main regulatory inputs that control protein kinase C activity are
- phosphorylation signals
- Ca^{2+} signals
- diacylglycerol signals
- binding to scaffolding or anchoring proteins.

Functions and regulation of protein kinase C are shown schematically in Fig. 7.11.

Regulation by Phosphorylation

Phosphorylation on Ser/Thr and Tyr residues is an essential step in the maturation and activation of PKC. An initial activating phosphorylation takes place on Ser/Thr residues in the activation segment of the catalytic domain. The reaction is catalyzed by the phosphoinositide-dependent protein kinase 1 (PDK1, see Section 6.6.3) and appears to occur already on newly synthesized enzyme. As a consequence of the PDK1-catalyzed phosphorylation, autophosphorylation on two sites near the C-terminus is triggered. One of the autophosphorylation sites is located in a hydrophobic segment, and its phosphorylation has been found to induce membrane dissociation of the activated enzyme. Overall, the PDK1-catalyzed phosphorylation and the autophosphorylation events are important for regulating the catalytic activity of the PKCs and for their subcellular distribution. The details of this regulation mechanism and its relation to activation by the other cofactors are not yet established.

A pivotal role in PKC activation is attributed to the PDK1 kinase, which phosphorylates and activates many of the PKC isoenzymes as well as other central protein kinases like the Akt kinase and the p^{60}S6 kinase (review: Dempsey et al., 2001).

Tyrosine phosphorylation is another regulatory modification of PKC enzymes. In response to a variety of stimuli, e. g., treatment with H_2O_2 (Konishi et al., 2001), a phosphorylation on tyrosine residues located near the C-terminus and between the catalytic domain and the regulatory domain has been observed. The Src and Lyn nonreceptor protein kinases (see Section 8.3) have been identified as the tyrosine kinases responsible for this phosphorylation.

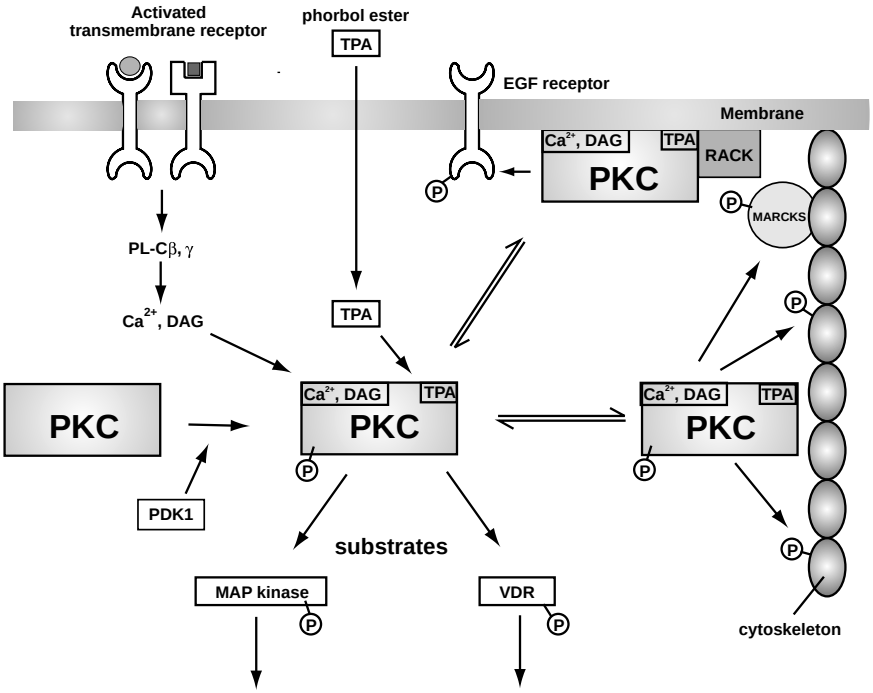


Fig. 7.11 Functions and regulation of protein kinase C. Receptor-controlled signal pathways lead to formation of the intracellular messenger substances Ca^{2+} and diacylglycerol (DAG) that, like phorbol ester (TPA), activate protein kinase C (PKC). For initial activation of PKC, phosphorylation by the protein kinase PDK1 is required. Translocation to the cell membrane is linked to further activation of protein kinase C; receptors for protein kinase C, the RACK proteins, are also involved. Substrates of protein kinase C are the MARCKS proteins and other proteins associated with the cytoskeleton. Other substrates are the Raf kinase (see Chapter 10) and the receptor for vitamin D_3 (VDR, see Chapter 4).

Regulation by membrane association

Many functions of protein kinase C in signaling pathways are closely linked with the membrane association of the enzyme. Signal-dependent association of PKC with the cell membrane is therefore another major regulatory aspect of this enzyme class. Activation of protein kinase C, initiated by addition of phorbol esters, for example, is associated with a redistribution of the enzyme from the cytosol to the membrane.

Membrane association, and thus activation of protein kinase C, is mainly controlled by the cofactors Ca^{2+}, diacylglycerol, and phosphatidyl serine, which enhance membrane association. Both the Ca^{2+} and the diacylglycerol signals can be generated by activation of phospholipase C isoenzymes (see Section 6.4), which in turn can be activated via two central signaling pathways, namely G protein-dependent pathways (see Section 5.62 5) and receptor tyrosine kinase pathways (see Section 8.1.4). Binding of Ca^{2+} to the C2 domain leads to an increased membrane association and to activation by release of the catalytic center from interaction with the autoinhibitory structural ele-

ment. Further activation takes place by binding of diacylglycerol to the C1 domain, which directly favors membrane association and enhances binding of phosphatidyl serine to the C2 domain, whereby these ligands serve as an anchor for membrane association. Use of the two membrane-targeting domains C1 and C2 apparently helps to ensure high affinity, specificity, and regulation of the membrane interaction.

The importance of coordination of the Ca^{2+} and diacylglycerol signals is very nicely illustrated for protein kinase Cγ of brain (Oancea and Meyer, 1998). For persistent activation of protein kinase Cγ, it is necessary that high-frequency Ca^{2+} signals and a diacylglycerol signal should be active simultaneously. Low-frequency Ca^{2+} signals, in contrast, only lead to low activation in the presence of diacylglycerol. Because of this property, the function of a molecular device able to decode Ca^{2+} and diacylglycerol signals is assigned to protein kinase Cγ.

Regulation by Localization and Protein – Protein Interactions

A major regulatory aspect of proteinkinase C enzymes is the localization to distinct subcellular sites mediated by distinct protein-protein interactions. The PKC enzymes have been found to interact – in addition to substrate proteins – with a multitude of proteins of distinct functions, including proteins with scaffolding, adaptor or membrane-targeting function and inhibitory proteins. Examples of membrane-targeting proteins that specifically interact with protein kinase C enzymes are the receptors for activated C kinase, the RACK proteins, which specifically bind activated forms of protein kinase C. This class of proteins anchor PKC enzymes to the membrane, providing access to membrane-localized substrates (see Section 7.7). Furthermore, subtypes of the RACK proteins are involved in intracellular trafficking of PLC enzymes. Most of the PKC-RACK interaction is mediated by the C2 domain of the PKC enzymes.

Members of the A kinase-anchoring protein family (AKAPs) have also been identified as binding partners of PKC enzymes. As an example, AKAP79 assembles PKC into protein signaling complexes, keeping it in an inactive state. Upon receipt of Ca^{2+}/diacylglycerol signals, the PKC is released from the inhibitory complex. Specific binding and clustering of PKC enzymes have also been reported for the adaptor proteins Ina D (see Section 8.5) and for the adaptor protein caveolin, which targets signaling proteins to specific plasma membrane microdomains, the caveolae.

7.4.4 Functions and Substrates of Protein Kinase C

The members of the protein kinase C family are central signal proteins and, as such, are involved in the regulation of a multitude of cellular processes. A large number of substrates have been identified for which each PKC isoenzyme has a distinct substrate specificity. Of the many substrates of protein kinase C isoenzymes, the MARCKS proteins are highlighted as very well characterized and specific substrates of protein kinase C (review: Arbuzova et al., 2002). The abbreviation MARCKS stands for myristoylated, alanine-rich C-kinase substrate.

The MARCKS proteins are a family of proteins that are involved in physiologically important processes such as cell mobility, secretion, membrane transport, and in the regulation of the cell cycle. All these processes are associated with changes and restructuring of the actin cytoskeleton. The role of converting extracellular signals into changes in the structure of the actin cytoskeleton is attributed to the MARCKS proteins. A protein kinase C-mediated phosphorylation of the MARCKS proteins is involved in this regulation. These are widespread and specific substrates of protein kinase C, and their phosphorylation is used as an indicator of the activation of protein kinase C *in vivo*.

The MARCKS proteins are acidic proteins with a high content of the amino acids Ala, Gly, Pro and Glu. An N-terminal domain carries a lipid anchor in the form of myristinic acid, from which it is assumed that it mediates the association with the membrane. A basic effector domain is important for regulation of the MARCKS proteins; a binding site for Ca^{2+}/calmodulin and the phosphorylation site for protein kinase C are located in this domain (see Fig. 7.12).

In the unphosphorylated form and in the absence of Ca^{2+}, the MARCKS proteins bind to actin filaments and bring about crosslinking of the latter. Binding of Ca^{2+}/calmodulin or phosphorylation by protein kinase C inhibits the crosslinking activity. The MARCKS proteins can thus modulate the aggregation status of actin filaments and function as effectors for the conversion of extracellular signals that are carried into the cell via G protein-coupled receptors and/or tyrosine kinase receptors.

Further examples of substrates of protein kinase C are the epidermal growth factor receptor (see Chapter 8), an Na^+/H^+ exchanger protein, Raf kinase (Chapter 9) and N-methyl-D-aspartate (NMDA) receptors. Activation of protein kinase C may, as the examples show, act on other central signal transduction pathways of the cell; it may have a regulating activity on transcription processes, and it is involved in the regulation of transport processes and in neuronal communication. Many substrates of protein kinase C are membrane proteins, and it is evident that membrane association of protein kinase C is of great importance for the phosphorylation of these proteins.

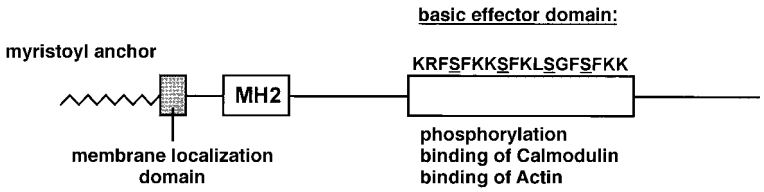


Fig. 7.12 Functional domains or the MARCKS proteins. Linear representation of the characteristic domains of the MARCKS proteins. The Ser phosphorylation sites in the effector domain are underlined. The function of the MH2 domain is unknown.

7.5
Ca^{2+}/Calmodulin-dependent Protein Kinases

7.5.1
Importance and General Function

The signal-mediating function of Ca^{2+} is performed as a Ca^{2+}/calmodulin complex in many signaling pathways. Ca^{2+}/calmodulin can bind specifically to effector proteins and modulate their activity. In first place as effector proteins of Ca^{2+}/calmodulin are the *Ca^{2+}/calmodulin protein kinases* (*CaM kinases*) (review: Fujisawa , 2001). The CaM kinases are widespread and are found in practically all cells of mammals. Like other target proteins of intracellular messengers, the CaM kinases also show great heterogeneity caused by the existence of distinct genes and alternative splicing.

A rough categorization of the CaM kinases differentiates between *specialized* CaM kinases and *multifunctional* kinases.

An example of a specialized CaM kinase is myosin light-chain kinase (MLCK), the primary function of which is to phosphorylate the light chain of myosin and thus to control the contraction of smooth musculature.

The multifunctional CaM kinases include the CaM kinases of types I, II and IV, all of which phosphorylate a rather broad spectrum of substrate proteins. These enzymes regulate many processes (see Table 7.1) such as glycogen metabolism, activity of transcription factors, microfilament formation, synaptic release of neurotransmitters from storage vesicles, biosynthesis of neurotransmitters, and many more. Some important substrates of CaM kinases are shown in Table 7.1. A primary function in synaptic transmission is ascribed to the CaM kinases of type II, which exhibit outstanding regulatory properties and will therefore be discussed in below in more detail. Among the class of CaM kinase II enzymes, subtypes α, β, γ and δ are differentiated. The α and

Tab. 7.1 Examples of substrates of Ca^{2+}/calmodulin dependent protein kinases. Source: Ann. Rev. Physiol. (1995), 57.

protein	function
AcetylCoA carboxylase	Biosynthesis of fatty acids
Glycogen synthase	Glycogen synthesis
HMGCoA reductase	Biosynthesis of cholesterol
NO synthase	Biosynthesis of NO
Ca^{2+} channel (N-type)	Presynaptic Ca^{2+} influx
Ca^{2+} ATPase (heart)	Storage of Ca^{2+}
Synaptogamin	Release of neurotransmitters
Ryanodin receptor	Release of Ca^{2+}
p56 Lck tyrosine kinase	Activation of T cells
EGF receptor	Growth control
Cyclic nucleotide phosphodiesterase	cAMP and CGMP metabolism
Phospholipase A2	Hydrolysis of phospholipids
Ribosomal protein S6	Protein biosynthesis
CRE binding protein	Transcription control

β subtypes of CaM kinase II only occur in the brain, whereas the other subtypes are also found in other organs. In the hippocampus, CaM kinase II constitutes up to 2 % of total cellular protein.

7.5.2 Structure and Autoregulation of CaM Kinase II

From a regulatory point of view, CaM kinase II is of particular interest, as it has the characteristic of an enzyme with a built-in "memory switch". The "memory" allows the CaM kinase II to conserve a stimulatory signal over a longer period of time and to remain in an activated state, even when the initiating stimulus has died away.

CaM kinase II is regulated by both autophosphorylation and Ca^{2+}/calmodulin (review: Fujisawa, 2001). An association domain, an N-terminal catalytic domain and a regulatory domain containing the Ca^{2+}/calmodulin binding site and an autoinhibitory region can be differentiated in the structure of CaM kinase II. Electron microscopic pictures show that CaM kinase II has an oligomeric structure (Fig. 7.13) in which 12 copies of the monomeric enzyme (α, β, γ or δ subtype) are configured in the form of a cylinder containing two stacks of 6 subunits (Kolodziej et al., 2001). The catalytic domains are orientated outwards, and the association domains form the central core of each ring. The aggregated form of CaM kinase is found as a homooligomer or as a heterooligomer composed of different subtypes. It represents the holoenzyme form, which performs the catalytic functions and is subject to a sophisticated control.

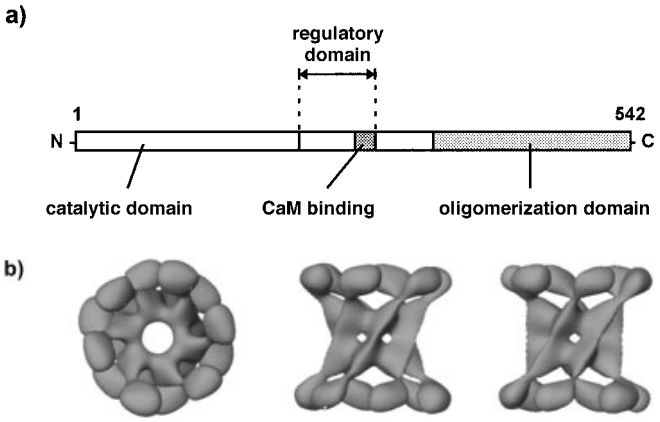


Fig. 7.13 a) Linear representation of the functional domains of CaM kinase II of type ß.
b) Oligomeric structure of CaM kinase II as determined from three-dimensional electron microscopy. The structure consists of 12 subunits that are organized in two stacked hexameric rings. The core of the structure is formed by the oligomerization domain whereas the outwards protruding parts are assumed to contain the catalytic domain and the calmodulin-binding domain.

In the absence of Ca^{2+}/calmodulin, the catalytic domain of CaM kinase II exists in an autoinhibited state. The active center is blocked by occupation by a pseudosubstrate, an autoinhibitory sequence of the regulatory domain. An increase in the Ca^{2+} concentration and the associated Ca^{2+} signal lead to binding of Ca^{2+}/calmodulin at the C-terminal end of the regulatory domain, which releases the enzyme from its inhibited state

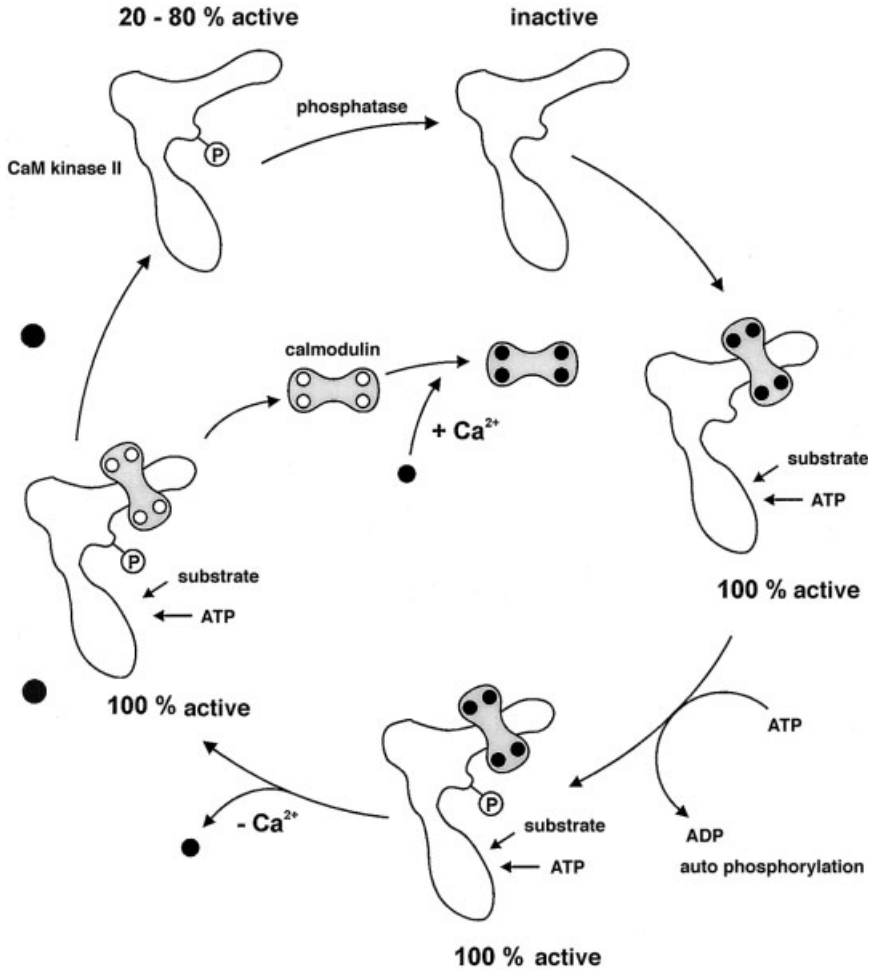


Fig. 7.14 Regulation of CaM kinase II. Scheme of regulation of CaM kinase II by Ca^{2+}/calmodulin and by autophosphorylation. CaM kinase II is inactive in the unphosphorylated form and in the absence of Ca^{2+}/calmodulin. Binding of Ca^{2+}/calmodulin activates the kinase for phosphorylation of protein substrates. In the process, autophosphorylation takes place at a conserved Thr residue that stabilizes the active state of the enzyme. In this state, significant residual activity is still present after dissociation of Ca^{2+}/calmodulin and the enzyme remains in an active state for a longer time after the Ca^{2+} signal has died away. The active state is only terminated when the activating phosphate residue is cleaved off by a protein phosphatase.

(Fig. 7.14). The enzyme is now activated and can perform an autophosphorylation in the autoinhibitory sequence. The phosphorylation takes place at a conserved Thr residue (Thr286 of the *α* subtype) and is intermolecular, i. e., neighboring subunits of the holoenzyme mutually phosphorylate one another.

The autophosphorylation has two important consequences:

- The affinity for Ca^{2+}/calmodulin is increased by close to three orders of magnitude. Ca^{2+}/calmodulin only dissociates very slowly from this high-affinity complex. The activated state is thus preserved over a longer period of time. Even when the Ca^{2+} signal has died away and the Ca^{2+} concentration has fallen to a level of 10^{-7} M, the enzyme remains in the activated state for several more seconds, since calmodulin can remain bound to the enzyme without bound Ca^{2+}.
- After dissociation of calmodulin, the phosphorylated enzyme still has 20–80 % of the activity of the Ca^{2+}/calmodulin-bound form. This ensures that significant activity remains after the Ca^{2+}/calmodulin signal has died away. In the phosphorylated form, CaM kinase II is in an autonomous, Ca^{2+}/calmodulin-independent state. This is only terminated when phosphatases cleave off the activating phosphate residue and thus lead the enzyme back into the inactive state.

The special feature of regulation of CaM kinase II is the memory effect within the activation process. Activation of the enzyme is initiated by a generally transient increase in cellular Ca^{2+}. Ca^{2+} activates CaM kinase II in the form of the Ca^{2+}/calmodulin complex; the kinase remains active even after the Ca^{2+} signal has died away, because the enzyme is converted into an autonomous activated state upon autophosphorylation. Presumably the complex holoenzyme structure of CaM kinase II endows the kinase with this unique regulatory property, allowing it to function as a sensor of cellular Ca^{2+} oscillations.

A special importance is attributed to this property, particularly for the detection and differentiation of repetitive Ca^{2+} signals in neuronal cells. The magnitude of constitutive CaM kinase II activity in the oligomeric, autophosphorylated form has been shown to depend on the duration, amplitude and frequency of elevated Ca^{2+}. For example, the interval between the occurrence of staggered Ca^{2+} signals is a determining factor for the intensity of activation. If the Ca^{2+} signals occur with a higher frequency, a long-lasting and effective activation is possible, since the kinase remains in the activated state between signals because of the memory effect (De Koninck and Schulman, 1998). All these regulatory possibilities appear to be based on the specific oligomeric holoenzyme structure. Because of these special properties, it is assumed that CaM kinase II actively participates in synaptic plasticity and memory formation. The ability of CaM kinase II to decode the frequency of Ca^{2+} oscillations during synaptic stimulation and to give a prolonged response beyond the initial stimulus enables it to provide two characteristics required for a molecule involved in synaptic plasticity and in memory formation. In agreement with this is the observation that the memory capability of transgenic animals is influenced by expression of CaM kinase II (Mayford et al., 1996).

There are different mechanisms that lead to increased intracellular Ca^{2+} concentration and thus to activation of CaM kinase II (see Section 6.5 and Fig. 6.6). CaM kinases

of type II are activated as a consequence of $InsP_3$-mediated release of Ca^{2+} from intracellular storage. Influx of Ca^{2+} from the extracellular region, triggered by opening of various ligand-controlled or voltage-controlled Ca^{2+} channels, also brings about an activation of CaM kinase II. In addition to the Ca^{2+}/calmodulin signal, there are other proteins and enzymatic activities that are of importance for the efficiency and specificity of CaM kinase II signaling. One control is exerted by the activity of protein phosphatases that terminate the autoactivated state of CaM kinase II. Dephosphorylation of CaM kinase II is catalyzed by protein phosphatases I (see below) and by a specialized CaM kinase phosphatase. Another major regulatory aspect of CaM kinase II is its colocalization with substrates at specific subcellular sites. Subtypes of CaM kinase II have been found to be specifically localized in the nucleus and at the cytoskeleton, especially at postsynaptic structures called postsynaptic densities, PSD, which contain large amounts of associated CaM kinase II. By interaction with anchoring proteins like PDZ proteins (see Section 8.5), CaM kinase II is assembled in the vicinity of sites of Ca^{2+} entry and of potential substrates. Examples of neuronal substrates are the ligand-gated ion channels like the AMPA receptor, the NMDA receptor, and neuronal NO synthase (review: Fukunaga and Miyamoto, 2000).

The two other multifunctional CaM kinase of types I and IV are, apart from Ca^{2+}/calmodulin control, regulated via phosphorylation by an upstream CaM kinase kinase, CaMKK (review: Soderling, 1999). The CaMKK phosphorylates CaM kinase I and IV in the activation segment and thereby greatly enhances the activity of these enzymes. Examples of substrates of CaM kinases I and IV are transcription factors, the MAP kinases and adenylyl cyclase.

7.6 Ser/Thr-specific Protein Phosphatases

Under physiological conditions, phosphate esters of Ser and Thr residues are stable and show only a low rate of spontaneous hydrolysis. Thus, the cell requires its own tools for regulated cleavage of phosphate residues to terminate and damp signals mediated by protein phosphorylation. This role is performed by specific protein phosphatases.

7.6.1 Structure and Classification of Ser/Thr Protein Phosphatases

We know of Ser/Thr-phosphate-specific protein phosphatases and Tyr-phosphate-specific protein phosphatases. The latter are dealt with in Chapter 8.

In mammals, at least eight types of Ser/Thr-specific protein phosphatases have been identified, and these are classified according to their substrate specificities, metal requirements, and sensitivities to natural or synthetic inhibitors. Four groups of the Ser/Thr phosphatases, the *protein phosphatases (PP) 1, 2A, 2B* and *C2*, have been studied in detail. PP-1, PP-2A and PP-2B, which is also known as *calcineurin*, are highly homo-

logous with respect to the sequence of the catalytic domain, whereas PP2C appears to have a distinct evolutionary background. Generally, the protein phosphatases are made up of a catalytic subunit with one or more associated subunits that have regulatory and/or targeting functions.

7.6.2 Regulation of Ser/Thr Protein Phosphatases

Protein phosphatases are the antagonists of protein kinases. They perform a dual function. By diminishing and terminating a signal created by protein phosphorylation, they can have a damping effect on protein kinase-mediated signal transduction. Protein phosphatases can also have a positive, reinforcing effect in signaling pathways. Dephosphorylation of a signal protein by a protein phosphatase can lead to its activation and thus to amplification of the signal (Fig. 7.15).

Because of these functions, the protein phosphatases are an indispensable part of signal transduction processes involving protein phosphorylation. It is therefore not surprising that the protein phosphates are subject to diverse and complex regulation. A large part of this regulation is exerted via additional subunits that associate with the catalytic subunit to form the active holoenzyme. These subunits serve to transmit incoming signals to the catalytic subunit and to target the catalytic subunit to distinct subcellular sites.

Regulation of the Ser/Thr phosphatases takes place predominantly by the following mechanisms:

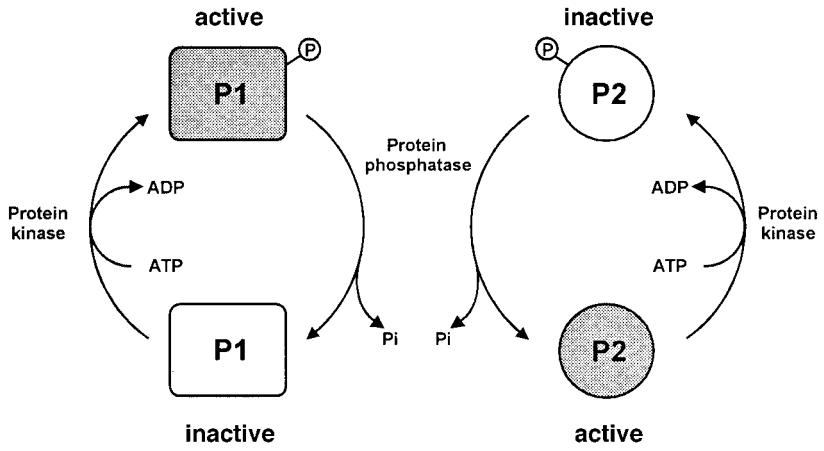


Fig. 7.15 The dual function of protein kinases and protein phosphatases. Phosphorylation of proteins (P1, P2) can fix the latter into an active or inactive state. In the case of P1, protein kinases have an activating effect and protein phosphatases are inactivating; the reverse is true for P2.

- **Phosphorylation**
  Targets of phosphorylation are both the catalytic and the regulatory/localization subunits. These phosphorylation events can change subunit composition, catalytic activity, and subcellular localization.
- **Binding of cofactors**
  For one subtype of PP2A, specific activation by the second messenger substance ceramide has been clearly established.
- **Targeted localization**
  The intrinsic substrate specificity of protein phosphatases, when measured *in vitro*, is quite low. Much of the specificity of substrate dephosphorylation is achieved by targeting or scaffolding subunits that serve to localize the phosphatase in proximity to particular substrates, and also to reduce its activity towards other potential substrates.

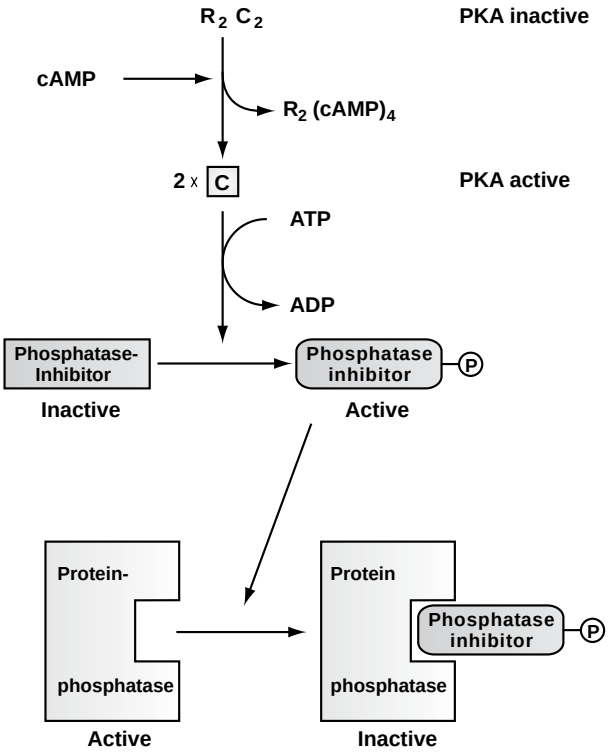


Fig. 7.16 Regulation of protein phosphatases by inhibitor proteins. The substrates of protein kinase A include protein phosphatase inhibitors that are phosphorylated by the C subunit of protein kinase A. In the phosphorylated state, the protein phosphatase inhibitors bind to the protein phosphatase and inhibit its enzyme activity.

- **Specific inhibitor proteins**
  Specific inhibitor proteins for Ser/Thr phosphatases exist which can control the activity of the protein phosphatases. These inhibitors are generally subject to regulation themselves, e. g., by phosphorylation (Fig. 7.16).

7.6.3 Protein Phosphatase I, PPI

PPI is a major class of eukaryotic Ser/Thr-specific protein phosphatases that regulate diverse cellular processes such as cell cycle progression, muscle contraction, carbohydrate metabolism, protein synthesis, transcription, and neuronal signaling. Its action is modulated and regulated by association with subunits including various inhibitor proteins and multiple targeting subunits of which nearly 30 proteins have now been identified (review: Aggen et al., 2000). The activity of the inhibitory proteins can be controlled via phosphorylation by protein kinase A as outlined in Fig. 7.16.

Of the many targeting subunits of PP1, the skeletal muscle glycogen binding subunit G_M has been best studied and will be presented below in more detail.

Regulation of PP1 in Glycogen Metabolism

Protein phosphatase I has an important role in the regulation of glycogen metabolism, where it is inhibited under conditions of glycogen degradation and activated under conditions of gylcogen synthesis. The basis of the regulation of protein phosphatase I in the skeletal muscle is the ability of the catalytic subunit of the enzyme to reversibly associate with a glycogen-binding protein, also known as the *G_M subunit of protein phosphatase I* (Fig. 7.17). As a result, the activity and specificity of the PP1 are modulated.

The G_M subunit is tightly bound to glycogen and localizes the PP1 to three glycogen-binding enzymes involved in glycogen metabolism of skeletal muscle: glycogen phosphorylase, phosphorylase kinase, and glycogen synthase. Association of the catalytic subunit with the G_M subunit creates a form of protein phosphatase I known as *protein phosphatase 1-G* (*PP1G*). PP1G is a highly active form of the protein phosphatase, in which the G_M subunit functions to mediate a targeted localization of the catalytic subunit to glycogen, so that the enzyme comes into the direct vicinity of its glycogen-bound substrates. When associated with G_M and glycogen, the PP1 efficiently dephosphorylates the phosphorylated forms of glycogen phosphorylase, phosphorylase kinase and glycogen synthase, and it is less sensitive to inhibition by cytosolic inhibitor proteins. Dephosphorylation of glycogen phosphorylase and phosphorylase kinase inactivates these enzymes. For activation of both enzymes, e. g., in response to an adrenalin signal, the PP1 must be removed from the glycogen. This is achieved by a phosphorylation of the G_M subunit. In response to adrenalin and activation of protein kinase A, G_M becomes phosphorylated in a cooperative way on two Ser residues, the P1 and P2 sites (Fig. 7.17A). Upon phosphorylation of both sites, the PP1 is released from the G_M subunit and migrates in from the glycogen stores to the cytosol, where it is deactivated by binding to inhibitor proteins. The overall effect is to promote

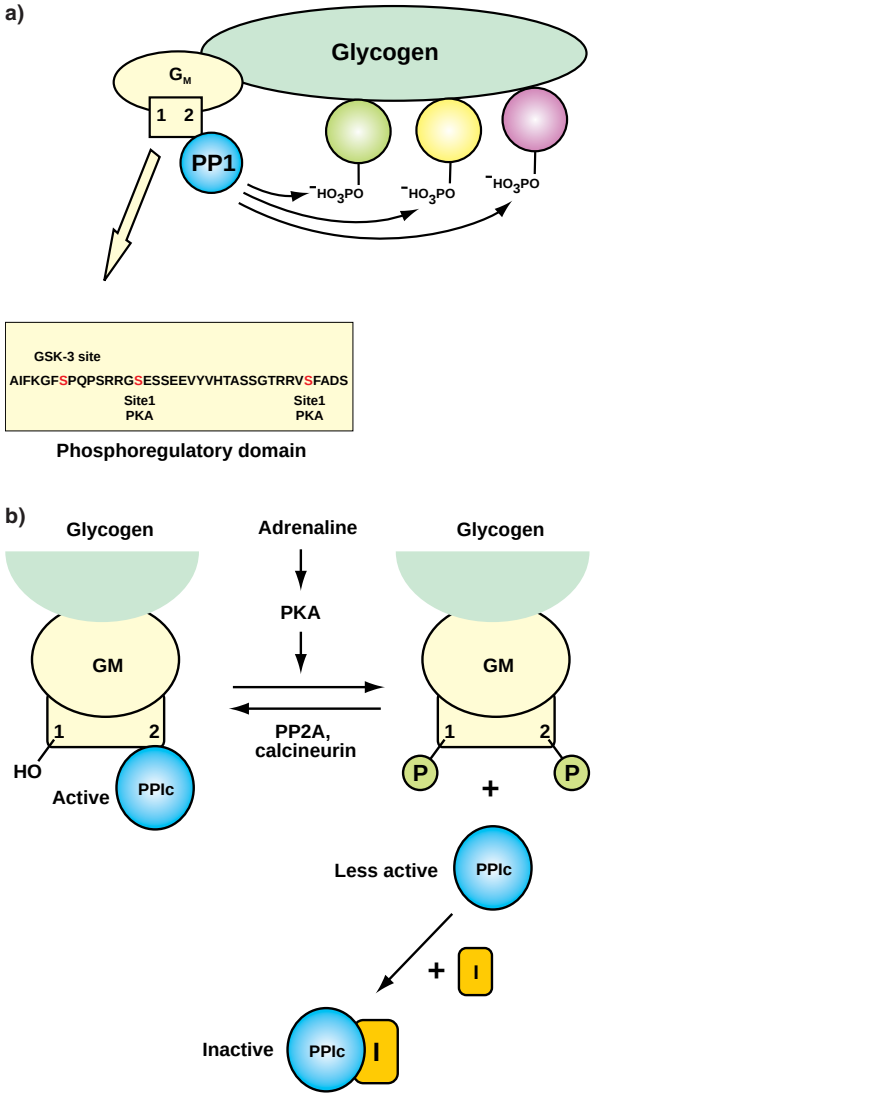


Fig. 7.17 a) Representation of the functional PP1-G holoenzyme bound to glycogen and juxtaposed with potential phosphorylated substrates.The boxed region of G_M is the region containing the phosphoregulatory sites 1 and 2, and the open arrow indicates an expansion of that region. The expansion shows the amino acid sequence of the phosphoregulatory domain and the PP1c- binding domain (site 2). Serine residues phosphorylated by cAMP-dependent protein kinase (PKA) and by Glycogen synthase-kinase-3 (GSK-3) kinase are indicated in red. The green, yellow and purple balls are phosphorylated PP1 substrates, such as phosphorylase, glycogen synthase and phosphorylase kinase.

b) Model for the phosphoregulation of PPI-G. The boxed region of G_M is the phosphoregulatory domain, and 1 and 2 refer to sites 1 and 2, respectively. The fully active PP1-G holoenzyme is shown in the upper left: Phosphorylation of sites 1 and 2 via protein kinase A induces dissociation of the catalytic subunit PP1c which is of low activity. Complexation of the liberated PP1c with the inhibitor I leads to ist inactivation.

glycogenolysis and mobilization of glucose from energy stores in response to the release of adrenalin.

Renewed docking of the catalytic subunit requires the removal of the phosphate residues at the G subunit. This takes place via the protein phosphatases 2A and 2B (calcineurin). Thus, a cascade of protein phosphatases is involved in the regulation of dephosphorylation of key enzymes of glycogen degradation, whereby a phosphatase, namely protein phosphatase I, is indirectly activated by other protein phosphatases. With calcineurin, a Ca^{2+}-dependent protein phosphatase is involved, and thus it is possible to influence glycogen metabolism via Ca^{2+}-mediated signals.

7.6.4
Protein Phosphatase 2A, PP2A

The family of protein phosphatase 2A enzymes accounts for most of the protein phosphatase activity in the cell. Contrary to the original belief that PP2A enzymes are only housekeeping enzymes, it is now increasingly recognized that PP2A activities are highly regulated and are involved in the control of numerous cellular processes including cell growth and proliferation, apoptosis, cell signaling, and intermediary metabolism (reviews: Janssens and Goris, 2001; Sontag, 2001). PPA2 enymes are oligomeric enzymes composed of a conserved catalytic subunit and one or more additional reg-

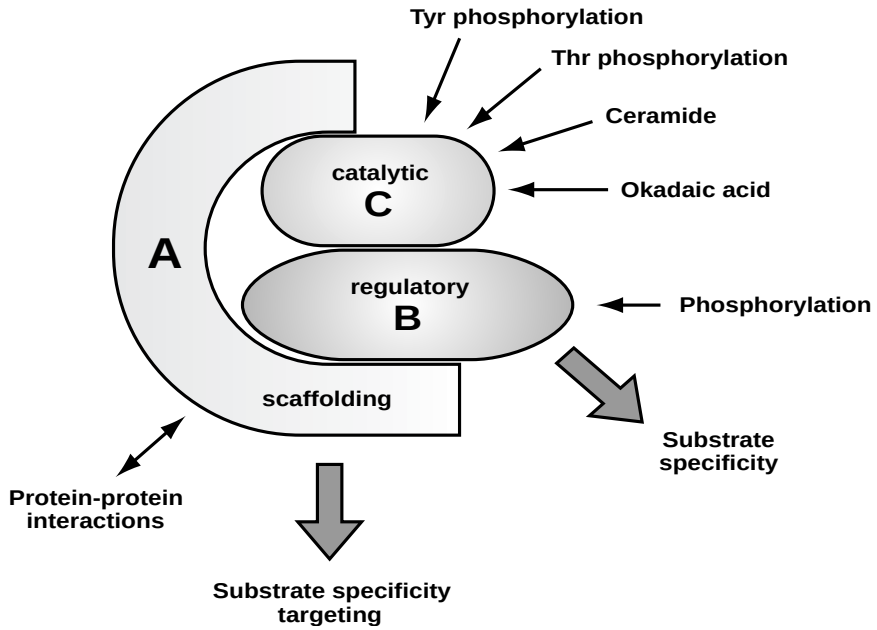


Fig. 7.18 Subunit structure of protein phosphatase 2A (PP2A). The common hetrotrimeric form of PP2A containing the catalytic subunit C, the regulatory subunit B and a structural subunit A is shown. In addition, various cellular proteins that interact with PP2A are indicated.

ulatory subunits (Fig. 7.18). The core PP2A enzyme is a dimer consisting of the catalytic subunit C and a regulatory A subunit of 65 kDa. A third B subunit of which four subfamilies with a total of more than 20 members are known can associate with the core enzyme to form the PP2A holoenzyme. Because of the existence of two isotypes of the A and C subunits and because of the large number of different B subunits, a multitude of trimeric PPA2 holoenzymes can form, each with distinct enzymatic and regulatory properties.

Regulation of PP2A activity is mediated mainly via modification of the C subunit and via the nature of the associated B and A subunits. Post-translational modification of the C subunit by phosphorylation of specific Thr and Tyr residues inactivates the enzyme. Interestingly, reactivation can occur by intramolecular autodephosphorylation. Tyr phosphorylation is believed to be catalyzed by receptor and nonreceptor tyrosine kinases like the EGF receptor, the insulin receptor kinase, and the nonreceptor tyrosine kinase $p56^{Lck}$ (see Section 8.3)), integrating PP2A activity into central signaling pathways. PP2A enzymes show both a Ser/Thr-phosphate-specific and a weaker Tyr-phosphate-specific protein-phosphatase activity. The latter activity can be specifically stimulated by a distinct protein named protein tyrosine phosphatase activator, PTPA.

A major determinant of PP2A activity is the identity of the associated B subunit. The nature of the B subunit assembled in the heterotrimer of PP2A influences substrate specificity, catalytic activity, and subcellular distribution of the holoenzyme. The striking features of the B subunits are their diversity, stemming from the existence of entire subunit families, and the apparent lack of sequence similarity between these gene families.

Other regulatory influences of PP2A activity include the binding of inhibitory proteins and the binding of the second messenger ceramide to specific holoenzyme forms.

7.6.5
Protein Phosphatase 2B, Calcineurin

Calcineurin (protein phosphatase 2B) is a Ser/Thr phosphatase that is controlled by cellular calcium and regulates a large number of biological responses including lymphocyte activation, neuronal and muscle development, and the development of vertebrate heart valves.

Like the other major classes of protein phosphatases, calcineurin is an oligomeric enzyme. It is made up of a catalytic subunit, calcineurin A, and a regulatory subunit, calcineurin B. In addition, binding of Ca^{2+}/calmodulin is required to form the fully active holoenzyme. The catalytic subunit harbors the catalytic domain and three regulatory domains, which have been identified as the calcineurin B binding site, the Ca^{2+}/calmodulin-binding domain, and an autoinhibitory domain. In the active site of calcineurin A, a binuclear metal center containing Fe^{2+} and Zn^{2+} is found. Both ions are thought to participate directly in phosphate ester hydrolysis (review: Rusnak and Mertz, 2001). Calcineurin B is highly homologous to calmodulin, containing 4 EF hands as binding sites for Ca^{2+}. In the absence of Ca^{2+}/calmodulin, the dimer of cal-

cineurin A and B is in an autoinhibited state caused by binding of the autoinhibitory domain to the substrate-binding cleft. Upon binding of Ca^{2+}/calmodulin, the autoinhibition is relieved. Both the Ca^{2+} binding to calcineurin B and the requirement of Ca^{2+}/calmodulin make calcineurin activity strongly Ca^{2+} dependent and allow it to function as a Ca^{2+} sensor.

The central regulatory function of calcineurin has been recognized during the search for the cellular target of the immunosuppressant drugs cyclosporin and FK506, often used in organ and tissue transplantations. The biochemical point of application of both pharmaceuticals was unclear for a long time. In initial experiments, it was found that cyclosporin and FK506 bind specifically to two proteins known as cyclophilin and FK506-binding protein, respectively. Both proteins function as peptidyl prolyl cis/trans isomerases. The immunosuppressive effect of cyclosporin and FK506 could not initially be explained by these observations. Only with the discovery that cyclosporin and FK506 achieve their immunosuppressive effect via inhibition of calcineurin did it become clear that the immunosuppression is mediated by a complex reaction chain involving calcineurin. It was shown that the complexes of cyclosporin/cyclophilin and FK506/FK506-binding protein bind to calcineurin and inhibit the phosphatase activity of the latter.

In the process of activation of T-lymphocytes, calcineurin is part of a signaling pathway that is activated by a rise in intracellular calcium upon ligand binding to the T-cell receptor (see Section 11.3) and finally leads to the activation of transcription factors, named NF-AT (Fig. 7.19). Members of the NF-AT family of transcription factors control the expression of a large number of proteins, including cytokines, ion channels, cell-surface proteins, and proteins involved in apoptosis. A subset of the NF-AT transcription factors (NF-ATc members) is found in the cytosol and undergoes nuclear translocation upon calcineurin activation, allowing subsequent activation of target genes (review: Crabtree, 2001). The cytoplasmic form of NF-ATc is phosphorylated in its nuclear localization signal and requires dephosphorylation by phosphatase action in order to get access to its cognate genes. At this point, Ca^{2+}-regulation of calcineurin comes into play.

The rise in Ca^{2+} caused by ligand binding to T cell receptors activates calcineurin's phosphatase activity, which dephosphorylates cytoplasmic NF-ATc proteins. Dephosphorylated NF-ATc enters the nucleus and binds to DNA in cooperation with other transcription factors, e. g., AP-1. In this way, many target genes in diverse tissues can be activated.

Transcription of interleukin 2 is inhibited as one of the consequences of inhibition of calcineurin by cyclosporin/cyclophilin and FK506/FK506-binding protein; as an extracellular signal, interleukin 2 can stimulate proliferation of lymphocytes. This discovery made it clear that the protein phosphatase calcineurin has an essential role in signal transduction in T lymphocytes and that these processes are severely impeded by cyclosporin and FK506.

Following the discovery of the role of calcineurin in transcription regulation in T-lymphocytes, a large number of calcineurin substrates other than NF-AT have been discovered. These include NO synthase, ion channels, and adenylyl cylase among many others (review: Rusnack and Mertz, 2000).

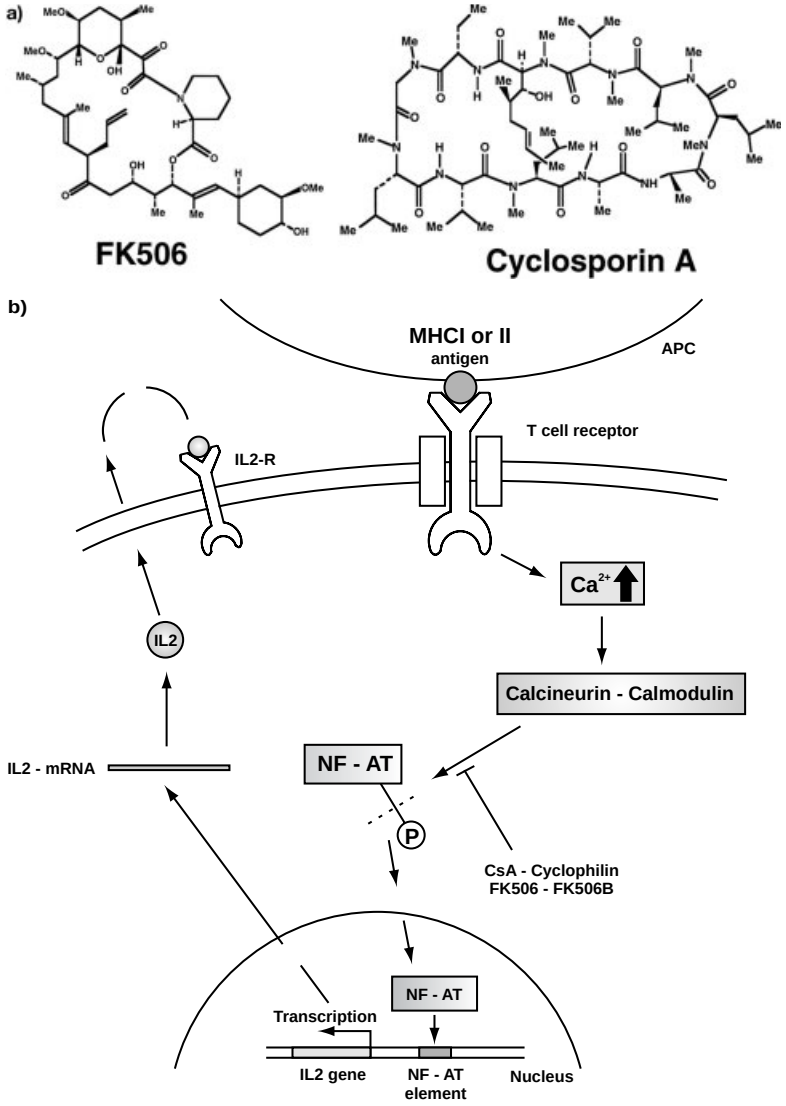


Fig. 7.19 a) Cyclosporin A and FK506. b) Model of the function of calcineurin in T lymphocytes. Antigenic peptides are presented to the T lymphocytes by an antigen-presenting cell (APC) within a cell-cell interaction (see also Chapter 11). Antigen binding activates the T cell receptor that starts a signal chain leading to an increase in cytosolic Ca^{2+} and activation of calcineurin. The activated calcineurin cleaves an inhibitory phosphate residue from the transcription factor NF-AT. Consequently, NF-AT is transported into the nucleus where it stimulates the transcription of corresponding genes. Amongst the genes controlled by NF-AT is the gene for the cytokine interleukin 2 (IL-2). Following secretion into the extracellular space, the IL-2 so formed binds to IL-2 receptors of the same cell or cells of the same type. A proliferation signal is created by the activated IL-2 receptor, leading to proliferation of T lymphocytes. Complexes of the immunosuppressants cyclosporin A (CsA) or FK506 with their binding proteins cyclophilin and FK506 binding protein (FK506B), respectively, inhibit calcineurin and disrupt the signal transmission to NF-AT.

7.7
Regulation of Protein Phosphorylation by Subcellular Localization

The extent and specificity of the reactions of protein kinases and protein phosphatases are extremely dependent on the degree to which substrate and enzyme are localized at the same place in the cell. Targeted localization of protein kinases and protein phosphatases is a major mechanism by which the selectivity of protein phosphorylation is enhanced and a tight control of phosphorylation, dephosphorylation, and interaction with cofactors is made possible. Many substrates of protein kinases occur either as membrane-associated or particle-associated forms. For protein kinases or protein phosphatases to perform their physiological function in a signal transduction process, they must in many cases be transported to the location of their substrate. This is valid both for the protein kinases as well as for protein phosphatases. In the course of activation of signal transduction pathways, compartmentalization of protein kinases redistributed to new subcellular locations is often observed. It is more and more recognized that the regulated assembly of protein phosphatases, protein kinases and their substrates into distinct multiprotein complexes – sometimes referred to as "transduceosomes" – is a major aspect of signaling by phosphorylation (Feliciello et al., 2001). By restricting the action of the two enzyme classes to a distinct subcellular site, the persistence, amplitude, and signal/noise ratio of phosphorylation signals are improved. Furthermore, signals from other effectors can be integrated more efficiently in these multiprotein signaling units.

Subcellular targeting of protein kinases and protein phosphatases is mediated by the regulatory subunits of these enzymes, which can bind specifically to scaffolding, adaptor or anchoring proteins located at distinct subcellular sites. Often these anchoring proteins are multivalent and allow the assembly of several signaling proteins. The mechanisms by which the anchoring proteins assemble at a distinct subcellular site are diverse. Structural membrane proteins, transmembrane receptors, ion channels, cytoskeletal proteins or cellular organelles can serve as the anchoring target. Other anchoring proteins interact with the membrane via lipid anchors or with distinct membrane anchoring sequences. Generally, the nature of the regulatory subunit of the protein kinase or protein phosphatase determines in which compartment of the cell and at which membrane section the protein phosphorylation signal will become active. The subunit functions as a *localization moiety*; it determines at which place in the cell the protein kinase or phosphatase gains access to its substrates.

The principle of targeted localization is shown in Fig. 7.20. In addition to the binding site for the corresponding protein kinase (or protein phosphatase), the regulatory subunit has a specific binding site for an *anchor protein*, found at a subcellular site in the region where protein phosphorylation should take place. Through the interaction of anchor protein and localization subunit, the catalytic subunit is fixed at the desired location and is able to preferentially convert substrate localized at the same location.

Among the protein kinases already discussed, there are some for which the function is linked with specific subcellular localization to a high degree. The anchoring proteins for protein kinases discussed below have been well characterized.

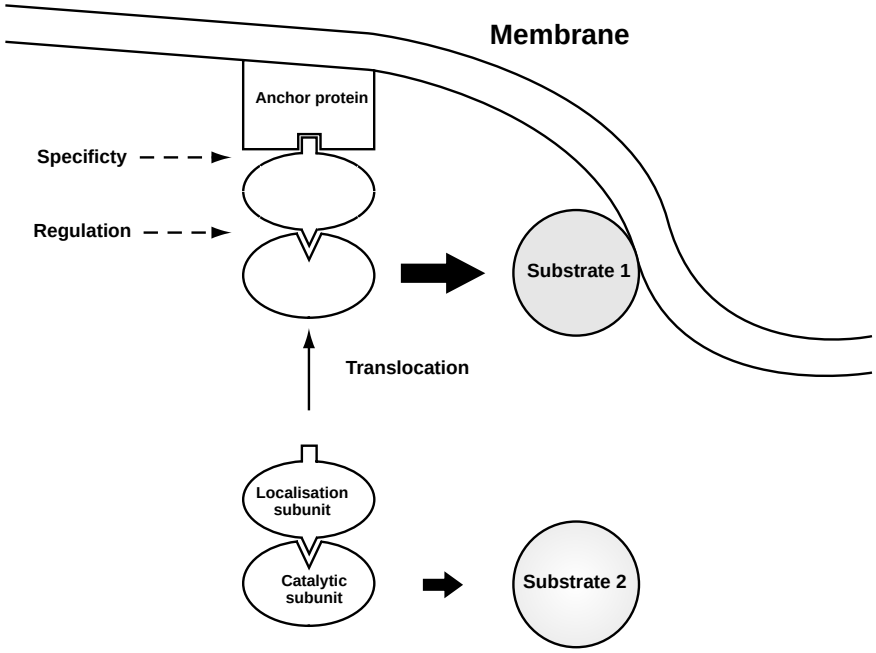


Fig. 7.20 The principle of targeted localization of protein kinases and protein phosphatases. The spatial configuration between the catalytic subunit of a protein kinase or protein phosphatase and a membrane-associated substrate is mediated by localization subunits that specifically bind to membrane-localized anchor proteins. The specificity of co-localization is predominantly achieved at the level of binding of the localization subunit to the anchor protein. The co-localization is regulated, in particular, by the interaction of the catalytic subunit with the localization subunit. In the membrane-associated form, the catalytic subunit has increased activity towards membrane-bound substrates.

A-Kinase Anchor Proteins, AKAPs

An increase in cAMP and activation of protein kinase A are accompanied, in many cases, by a change in the subcellular location of protein kinase A holoenzymes containing RII subunits (see Section 7.3). This targeted localization is mediated by the AKAPs, of which more than 20 different members are known (review: Feliciello et al., 2001). The AKAPs immobilize the PKA isoforms at specific subcellular sites by binding the RII subunits (Fig. 7.21). Binding of cAMP to the regulatory subunit releases the catalytic subunit that can phosphorylate substrates in the near vicinity. The released catalytic subunit can also be transferred to other compartments of the cell. Parallel to the increase in cAMP, translocation of the catalytic subunit is observed in many cells from the Golgi apparatus to the nucleus via the cytosol, and is accompanied by stimulation of transcription.

In addition to the RII subunit of protein kinase A, the AKAP proteins also bring protein phosphatases and other protein kinases to specific subcellular sites in a targeted fashion. The AKAP79 protein is a neuronal protein associated with the cytoskeleton that binds protein kinase C and protein phosphatase 2B (calcineurin) as well as

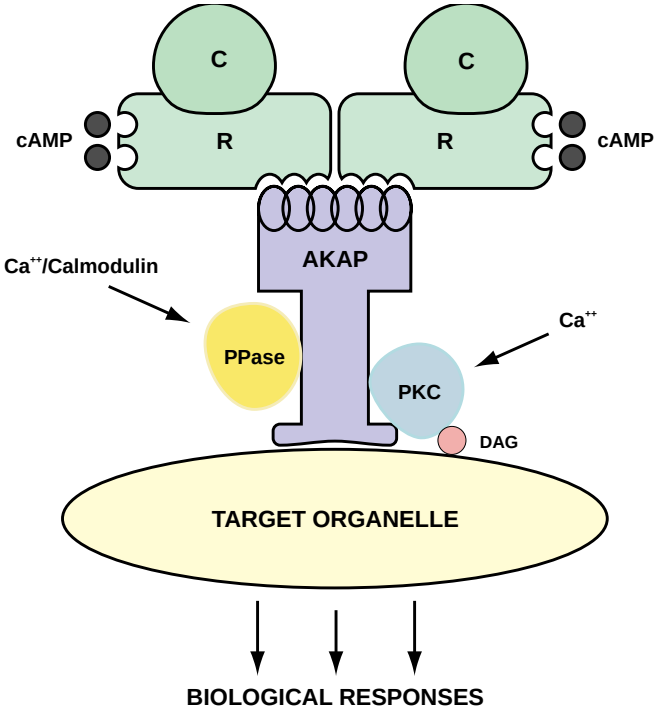


Fig. 7.21 Model of a signaling complex organized at the membrane of a target organelle by AKAPs. A schematic representation of a multiprotein complex formed by a prototypic AKAP, protein kinase A and two other effector proteins, protein phosphatase (PPase) and protein kinase C (PKC), both activated by calcium. DAG: diacylglycerol.

the RII subunit. The possibility of bringing a protein kinase and a protein phosphatase to the same place in the cell opens up the prospect of a coordinated and layered regulation of both enzyme activities in a multiprotein complex anchored at a distinct subcellular site. In this transduceosome, several signals can be integrated and directly transmitted to the substrates.

Specific association of AKAPs with ion channels and transmembrane receptors has also been observed. As an example, the AKAP named Yotiao directs both PKA II and protein phosphatase I to *N*-methyl-D-aspartate receptors.

By bringing together various protein kinases and protein phosphatases, anchoring proteins organize signal transduction events and can create localized and efficient signal events at specific subcellular sites.

Receptors for Protein Kinase C, RACK Proteins

The specific function of individual protein kinase C isoenzymes is controlled to a large extent by isoenzyme-selective adaptor proteins, collectively called "receptors for activated protein kinase C" or RACK proteins. In many cases, stimulation of cells with

phorbol esters or with hormones that activate phospholipase Cβ or phospholipase Cγ leads to translocation of protein kinase C isoenzymes from the cytoplasm to the cell membrane or cytoskeleton or into the nucleus. The differential localization of the protein kinase C isoenzymes is mediated by PKC-targeting proteins, of which the RACK proteins stand out (review: Schechtman and Mochly-Rosen, 2001). The RACK proteins belong to a class of targeting proteins that mediate anchoring of PKC to the membrane. These proteins bind to a structural section of protein kinase C enzymes that is not identical to the catalytic center or the substrate-binding site. Protein kinase C can thus be brought to the membrane in a targeted manner without any major intervention in catalytic activity. In addition to anchoring activated protein kinase C enzymes, the RACK proteins anchor other central signaling proteins. As an example, RACK 1, the anchoring protein for activated PKCβII, targets Src tyrosine kinase (see Section 8.3.2), cAMP-specific phosphodiesterase and integrin β-subunits (see Section 11.4) to distinct subcellular sites. RACK 1 contains several WD motifs that are thought to be responsible for the protein-protein interactions and for its scaffolding function. Similar to the AKAPs, the RACK proteins appear to be able to organize various signaling enzymes into multiprotein signaling complexes.

Reference

1 Aggen, J.B., Nairn, A.C. and Chamberlin, R. (2000) Regulation of protein phosphatase-1. *Chem.Biol.*, **7**, R13-R23.

2 Arbuzova, A., Schmitz, A.A. and Vergeres, G. (2002) Cross-talk unfolded: MARCKS proteins. *Biochem.J.*, **362**, 1–12.

3 Biondi, R.M. and Nebreda, A.R. (2003) Signalling specificity of Ser/Thr protein kinases through docking site-mediated interactions. *Biochem.J.*,

4 Crabtree, G.R. (2001) Calcium, calcineurin, and the control of transcription. *J.Biol.Chem.*, **276**, 2313–2316.

5 De Koninck, P. and Schulman, H. (1998) Sensitivity of CaM kinase II to the frequency of Ca2+ oscillations. *Science*, **279**, 227–230.

6 Dempsey, E.C., Newton, A.C., Mochly-Rosen, D., Fields, A.P., Reyland, M.E., Insel, P.A. and Messing, R.O. (2000) Protein kinase C isozymes and the regulation of diverse cell responses. *Am.J.Physiol Lung Cell Mol.Physiol*, **279**, L429-L438.

7 Engh, R.A. and Bossemeyer, D. (2001) The protein kinase activity modulation sites: mechanisms for cellular regulation - targets for therapeutic intervention. *Adv.Enzyme Regul.*, **41**, 121–149.

8 Feliciello, A., Gottesman, M.E. and Avvedimento, E.V. (2001) The biological functions of A-kinase anchor proteins. *J.Mol.Biol.*, **308**, 99–114.

9 Fujisawa H. (2001) Regulation of the activities of multifunctional Ca2+/calmodulin-dependent protein kinases. *J Biochem (Tokyo)*. **129**, 193–9.

10 Fukunaga, K. and Miyamoto, E. (2000) A working model of CaM kinase II activity in hippocampal long-term potentiation and memory. *Neurosci.Res.*, **38**, 3–17.

11 Hu, S.H., Parker, M.W., Lei, J.Y., Wilce, M.C., Benian, G.M. and Kemp, B.E. (1994) Insights into autoregulation from the crystal structure of twitchin kinase. *Nature*, **369**, 581–584.

12 Janssens, V. and Goris, J. (2001) Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem.J.*, **353**, 417–439.

13 Johnson, L.N. and O'Reilly, M. (1996) Control by phosphorylation. *Curr.Opin.Struct.Biol.*, **6**, 762–769.

14 Johnson, L.N., Noble, M.E. and Owen, D.J. (1996) Active and inactive protein kinases: structural basis for regulation. *Cell*, **85**, 149–158.

15 Johnson, L.N., Lowe, E.D., Noble, M.E. and Owen, D.J. (1998) The Eleventh Datta Lec-

ture. The structural basis for substrate recognition and control by protein kinases. *FEBS Lett.*, **430**, 1–11.

16 Kolodziej, S.J., Hudmon, A., Waxham, M.N. and Stoops, J.K. (2000) Three-dimensional reconstructions of calcium/calmodulin-dependent (CaM) kinase IIalpha and truncated CaM kinase IIalpha reveal a unique organization for its structural core and functional domains. *J.Biol.Chem.*, **275**, 14354–14359.

17 Konishi, H., Yamauchi, E., Taniguchi, H., Yamamoto, T., Matsuzaki, H., Takemura, Y., Ohmae, K., Kikkawa, U. and Nishizuka, Y. (2001) Phosphorylation sites of protein kinase C delta in H2O2-treated cells and its activation by tyrosine kinase in vitro. *Proc.Natl.Acad.Sci.U.S.A*, **98**, 6587–6592.

18 Mayford, M., Bach, M.E., Huang, Y.Y., Wang, L., Hawkins, R.D. and Kandel, E.R. (1996) Control of memory formation through regulated expression of a CaMKII transgene. *Science*, **274**, 1678–1683.

19 Mayr, B. and Montminy, M. (2001) Transcriptional regulation by the phosphorylation-dependent factor CREB. *Nat.Rev.Mol.Cell Biol.*, **2**, 599–609.

20 Mellor, H. and Parker, P.J. (1998) The extended protein kinase C superfamily. *Biochem.J.*, **332 (Pt 2)**, 281–292.

21 Oancea, E. and Meyer, T. (1998) Protein kinase C as a molecular machine for decoding calcium and diacylglycerol signals. *Cell*, **95**, 307–318.

22 Ron, D. and Kazanietz, M.G. (1999) New insights into the regulation of protein kinase C and novel phorbol ester receptors. *FASEB J.*, **13**, 1658–1676.

23 Runnels, L.W., Yue, L. and Clapham, D.E. (2001) TRP-PLIK, a bifunctional protein with kinase and ion channel activities. *Science*, **291**, 1043–1047.

24 Rusnak, F. and Mertz, P. (2000) Calcineurin: form and function. *Physiol Rev.*, **80**, 1483–1521.

25 Schechtman, D. and Mochly-Rosen, D. (2001) Adaptor proteins in protein kinase C-mediated signal transduction. *Oncogene*, **20**, 6339–6347.

26 Skalhegg, B.S. and Tasken, K. (2000) Specificity in the cAMP/PKA signaling pathway. Differential expression, regulation, and subcellular localization of subunits of PKA. *Front Biosci.*, **5** , D678-D693.

27 Soderling, T.R. (1999) The Ca-calmodulin-dependent protein kinase cascade. *Trends Biochem.Sci.*, **24**, 232–236.

28 Sontag, E. (2001) Protein phosphatase 2A: the Trojan Horse of cellular signaling. *Cell Signal.*, **13**, 7–16.

29 Zhang, G., Kazanietz, M.G., Blumberg, P.M. and Hurley, J.H. (1995) Crystal structure of the cys2 activator-binding domain of protein kinase C delta in complex with phorbol ester. *Cell*, **81** , 917–924.

8
Signal Transmission via Transmembrane Receptors with Tyrosine-Specific Protein Kinase Activity

One of the fundamental mechanisms by which multicellular organisms communicate is the binding of protein ligands to transmembrane receptors that possess tyrosine kinase activity, the receptor tyrosine kinases. Activation of tyrosine-specific protein kinase activity is triggered, in particular, by signals that control cell growth and differentiation. Extracellular signals are often protein hormones, which – if they have a regulating influence on cell proliferation – are also classed as growth factors. Table 8.1 shows a selection of mammalian growth factors, together with the corresponding receptors that conduct the signal into the cell interior. In addition to the receptor tyrosine kinases, there exists a large family of nonreceptor tyrosine kinases which are integral components of signaling cascades triggered by receptor tyrosine kinases and other transmembrane receptors.

Coupling of an extracellular signal to tyrosine phosphorylation in the interior of the cell can take place by two means and involves two different types of receptor (Fig. 8.1):

- **Transmembrane receptors with intrinsic tyrosine kinase activity**
  Some transmembrane receptors possess intrinsic tyrosine kinase activity. These receptors are known as *receptor tyrosine kinases*. Ligand binding to an extracellular domain of the receptor is coupled to the stimulation of tyrosine kinase activity localized on a cytoplasmic receptor domain. The ligand-binding domain and the tyrosine kinase domain are part of one and the same protein.
- **Transmembrane receptors with associated tyrosine kinase activity**
  Another type of transmembrane receptor is associated, on the cytoplasmic side, with a tyrosine kinase that is activated when a ligand binds to the extracellular receptor domain (see Chapter 11). The tyrosine kinase and the receptor are not located on the same protein in this case.

8.1
Structure and Function of Receptor Tyrosine Kinases

Receptor tyrosine kinases possess binding sites at the surface of the cell membrane that are specific for extracellular ligands. Ligand binding to the receptor activates a tyrosine-specific protein kinase activity of the receptor, located on the cytoplasmic

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

Tab. 8.1 Selected growth factors and growth factor receptors from mammals.

Growth factaor	Characteristics	Receptors
Platelet derived growth factor, PDGF, types AA, AB and BB	Dimers, A (17 kD)- and B (16 kD) chains, B chain ist product of *c-sis* protooncogene	2 types of receptor tyrosine kinases, PDGF-Rα (170 kD), PDGF-Rβ (180 kD)
Epidermal growth factor, EGF Transforming growth factor-α, TGF-α	ca 6 kD, EGF und TGF-α are up to 40% identical	Receptor tyrosine kinase. EGF-R is a product of the c-erbB proto-oncogene
Transforming growth factor-β, TGF-β1, -β2, -β3	Homodimer of 25 kD	TGFβ receptor I und II, contains Ser/Thr specific protein kinase activity
Insulin-like growth factor, IGF-1 and IGF-2	7kD, related to proinsulin	Receptor tyrosine kinase, IGF-R
Fibroblast growth factor, FGF-1, FGF-2, FGF-3, FGF-4	related protein of 16-32 kD	Receptor tyrosine kinase
Granulocyte colony stimulating factor, G-CSF	24 kD	150 kD, G-CSF-R, receptor with associated tyrosine kinase
Granulocyte macrophage-colony stimulating factor, GM-CSF	14 kD	51 kD, GM-CSF-R, receptor with associated tyrosine kinase
Interleukins 1-7, IL-1, Il-7, Interleukin 9, 12, 15, IL-9, IL-12, IL-15		IL-R-1 – Il-R-7, IL-R-9, IL-R-12, IL-R-15, receptors with associated tyrosine kinase
Interleukin 8, IL-8		IL-R-8, G-protein coupled receptor
Erythropoietin		Epo-R, receptor with associated tyrosine kinase
Tumor necrosis factor, TNF		TNF-R, receptor with associated tyrosine kinase
Leukemia inhibitory factor, LIF		LIF-R, receptor with associated tyrosine kinase
Interferon α, β, γ		INF-Rα, INF-Rβ, INF-Rγ, receptors with associated tyrosine kinase

domain. Consequently, tyrosine phosphorylation is initiated at the receptor itself and also on associated substrate proteins; these in turn trigger the biological response of the cell by switching on a further chain of reactions. The response can reach as far as the cell nucleus, where transcription of particular genes is activated. It can also affect the reorganization of the cytoskeleton, cell-cell interactions, and reactions of intermediary metabolism. In particular, the receptor tyrosine kinases regulate cell division activity, differentiation, and cell morphogenesis by this mechanism (review: Pawson, 2002).

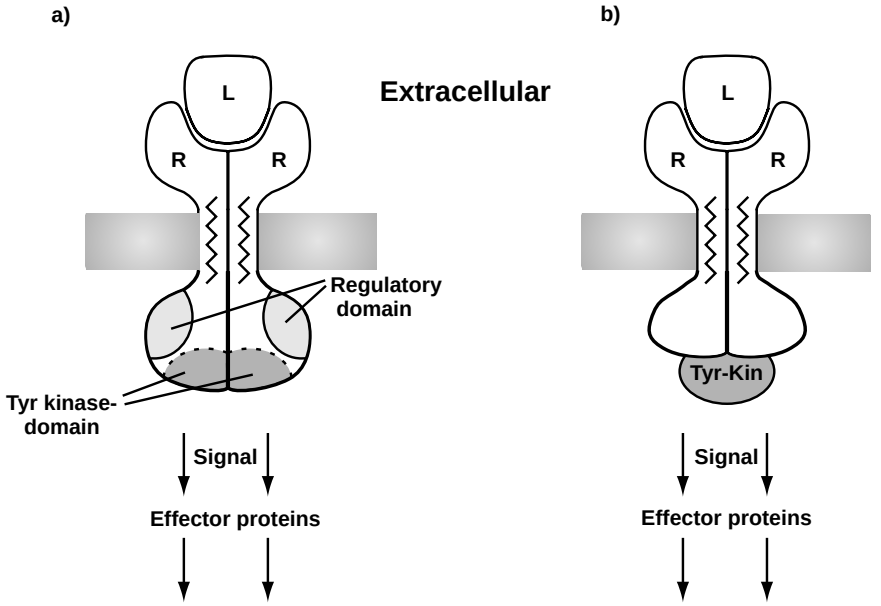


Fig. 8.1 Scheme of signal transmission by receptors with intrinsic and associated tyrosine kinase activity. a) Tyrosine kinase receptors possess a tyrosine kinase domain in the cytoplasmic region. Binding of a ligand L to the extracellular domain of the receptor produces a signal on the cytoplasmic side by activating the tyrosine kinase. Regulatory sequence segments are also located on the cytosolic side. b) Receptors with associated tyrosine activity pass the signal on to a tyrosine kinase that is not an intrinsic part of the receptor, but is permanently or transiently associated with the cytoplasmic receptor domain. The receptor shown has been simplified as a dimer.

8.1.1 General Structure and Classification

Receptor tyrosine kinases are integral membrane proteins that have a ligand-binding domain on the extracellular side and a tyrosine kinase domain on the cytosolic side (see Fig. 8.1). The transmembrane portion is made up of just one structural element; thus it is assumed that it crosses the membrane in an α-helical form. On the cytoplasmic side, in addition to the *conserved tyrosine kinase domain*, there are also further regulatory sequence portions at which autophosphorylation, and phosphorylation and dephosphorylation by other protein kinases and by protein phosphatases can take place.

The large family of mammalian receptor tyrosine kinases can be divided into different subfamilies, which are named according to their naturally occurring ligands. The subfamilies are classified according to the structure of the extracellular ligand-binding domains, in which different sequence portions can be differentiated (Fig. 8.2). In the extracellular domain, for example, there are Cys-rich sequences that occur as multiple repeats and sequences with an immunoglobulin-like structure. Most receptors are found as single polypeptide chains and are monomeric in the absence of ligand. One exception is the insulin receptor and its family members

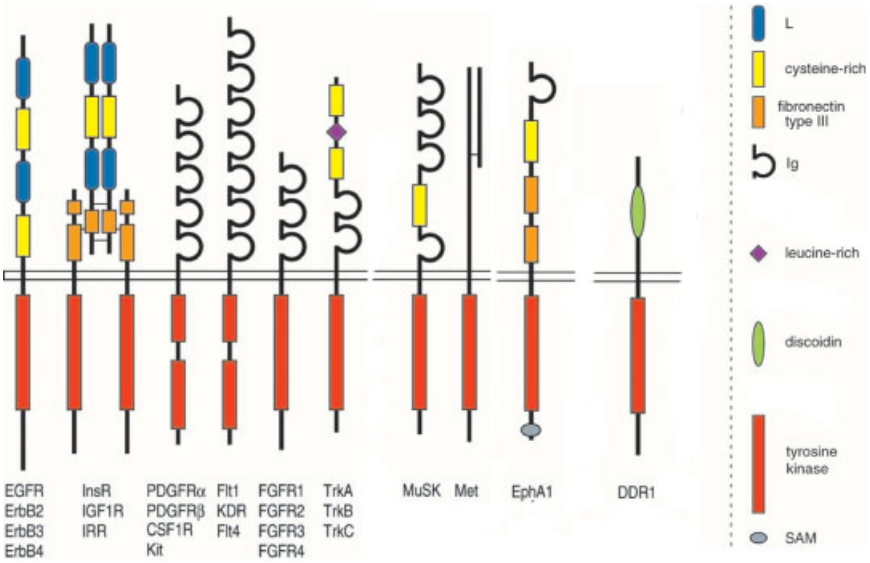


Fig. 8.2 Domain organization for a variety of receptor tyrosine kinases.
The extracellular portion of the receptors is on top and the cytoplasmic portion is on bottom. The lengths of the receptors is only approximately to scale.
Domains: L: homology domain; Ig: immunoglobulin-like domain;SAM: sterile alpha motif; Cadherin: cell-cell adhesion domain; Explanations to some of the receptors: PDGFR: platelet derived growth factor receptor; EGFR: epidermal growth factor receptor; InsR: insulin receptor; IGF1R: insulin-like growth factor 1 receptor; CSF1R: colony stimulating factor 1 receptor; KDR, Flt: receptors for the vascular endothelial growth factor; FGFR: fibroblast growth factor receptor; Trk: receptor for neurotrophins, such as nerve growth factor; MuSK: muscle specific receptor tyrosine kinase; Met: receptor for hepatocyte growth factor; Eph: receptor for ephrin ligands; DDR: discoidin domain receptor.

that comprise two extracellular α-chains and two membrane-spanning β-chains, forming a heterotetramer stabilized by various disulfide linkages.

8.1.2 Ligand Binding and Activation

The tyrosine kinase activity localized on the cytosolic side of the receptor is stimulated by extracellular ligand binding. Activation by extracellular ligands is based on a ligand-mediated *oligomerization* of the receptor. Ligand binding to the extracellular portions of the receptor tyrosine kinase induces the noncovalent oligomerization – mostly dimerization – of monomeric receptors, or it induces a structural rearrangement in a preassembled oligomeric receptor facilitation tyrosine autophosphorylation. Stimulation of the tyrosine kinase activity can have a twofold effect: Firstly, an autophosphorylation of the receptor may take place in trans, i. e., between different promotors of the receptor (Fig. 8.3). Secondly, substrate proteins may be phosphorylated on Tyr residues.

Ligand-mediated change of the dimeric structure of the receptor is a general method of transmission of signals through the membrane into the cell interior, with the help of

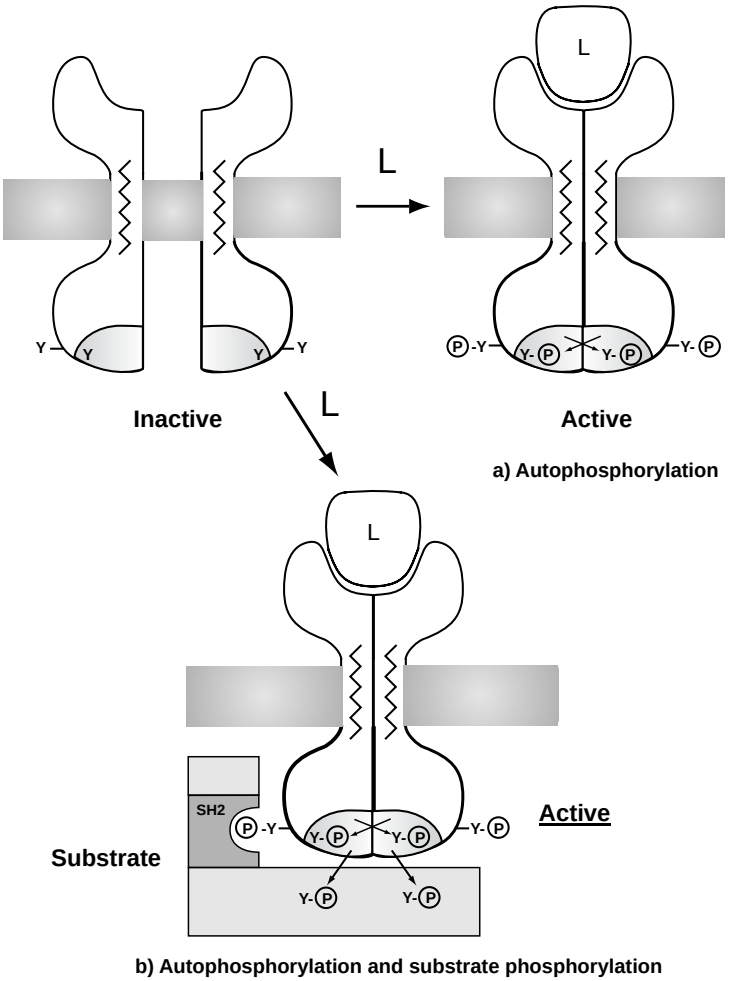


Fig. 8.3 Ligand-induced autophosphorylation and substrate phosphorylation of receptor tyrosine kinases. The tyrosine kinase domain of the receptor tyrosine kinase is activated by ligand binding. Consequently, autophosphorylation and/or phosphorylation of substrate proteins takes place. The substrate proteins possess specific phosphotyrosine binding domains (SH2 in the figure or PTP domains, see 8.2), which bind to phosphate residues formed in the process of autophosphorylation.

receptor tyrosine kinases (reviews: Heldin, 1995; Hubbard and Till, 2000). It is assumed that signal transmission via receptors with associated tyrosine kinases takes place by a similar mechanism. However, the situation is more complicated here because receptors with associated tyrosine kinase activity are often composed of many subunits (see Chapter 11).

There are at least two mechanisms by which ligand binding can change the oligomer status of the receptor (Fig. 8.4), by induction of dimerization (or higher oligomerization) or by induction of conformational changes in a preformed receptor oligomer:

For *ligand-induced dimerization*, two pathways have been described. In the first pathway, which applies for the growth hormone receptor and the erythropoietin receptor (see Chapter 12), the monomeric ligand has two binding sites for the receptor molecule and brings about a dimerization of the receptor. In the absence of the ligand, the

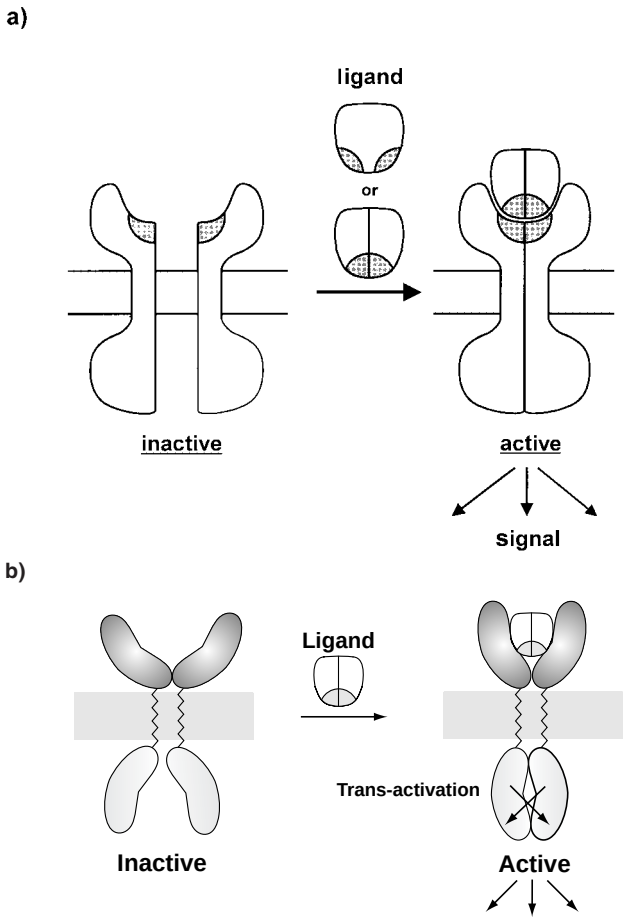


Fig. 8.4 Mechanism of activation of receptor tyrosine kinases by ligand binding. Activation of receptor tyrosine kinases is based on a ligand-induced oligomerization and/or conformational change of the receptor. An example is shown of a dimeric receptor; however, activation can also occur in a higher receptor oligomer. A) A bivalent ligand (monomer or dimer) induces a dimerization of a receptor which exists in a monomeric form without the ligand. B) A dimeric receptor is activated via an allosteric mechanism by ligand binding. In the absence of the ligand, the two kinase active sites are not close enough for mutual activation by phosphorylation (trans-activation).

receptor exists in a monomeric form. In another pathway, two ligand molecules are required to bring about dimerization of the receptor (see below, EGF).

Another mechanism of receptor activation uses *conformational changes within a preformed receptor dimer* (or higher oligomer). Here, the bound ligand appears to *stabilize* a distinct conformation of the *pre-assembled oligomeric form* of the receptor, fixing the receptor in a catalytically active state. The insulin receptor, for example, is a heterotetrameric protein composed of two $\alpha\beta$-units linked by disulfide bridges. It is assumed that in this case the activation takes place by an *allosteric mechanism*. The extracellular ligand binding brings about a change in the relative configuration of the two tyrosine kinase domains, in such a way that mutual Tyr phosphorylation is enabled. In the absence of the ligand, the two active sites are thought to be too distant for this trans-phosphorylation.

Ligand Structure and Receptor Oligomerization

The ligands of the receptor tyrosine kinases are generally multivalent. The multiple binding sites of the ligand enable it to orient two (or more) of the receptor subunits, so that their cytosolic tyrosine kinase activity is stimulated.

In some cases, the ligand itself has a dimeric structure and induces the formation of active receptor dimers on binding to the receptor. Fig. 8.5a shows the structure of the complex between the vascular endothelial growth factor (VEGF) and its receptor, FLt1. This structure represents the simplest dimerization scenario: a dimeric ligand engaging two receptors.

Other ligands, such as the epidermal growth factor (EGF), human growth hormone (hGH) and fibroblast growth factor (FGF), exist as monomeric proteins in solution. How monomeric ligands can induce a change in the oligomer structure of the receptor is illustrated by the example of hGH; a crystal structure is available of the complex of this ligand with the ligand-binding domain of the hGH receptor, which belongs to the class of cytokine receptors (see Chapter 11). The structure proves that the monomeric hGH protein is able to bind two molecules of the ligand-binding domain of the receptor (Fig. 8.5b). Accordingly, the monomeric hGH protein is divalent with respect to binding of its receptor.

In the case of the FGF receptor whose ligand FGF is monomeric in solution, another mechanism of ligand-induced receptor dimerization has been proposed. According to this model, the FGFR uses an additional ligand, namely heparin or heparan sulfate proteoglycans to stabilize the 2:2 complex between the FGF and the FGFR. The additional heparin ligand has been shown to bridge two receptor molecules in the dimer and the adjoining FGF ligands, allowing for stable dimerization of the ligand-bound receptor. Another example for this type of receptor activation is the EGF receptor. Here, a 1:1 receptor-ligand complex dimerizes via a specific loop located in a cysteine-reich region of the extracellular domain of the receptor. (review: Schlessinger, 2002).

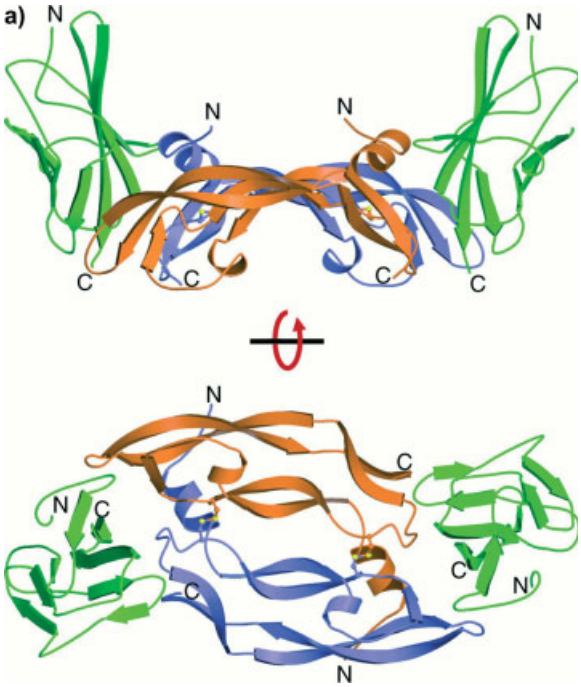


Fig. 8.5 a) Mechanism of dimerization of the vascular endothelial growth factor (VEGF) and its receptor Flt1. The dimeric ligand VEGF engages two receptor molecules. A ribbon diagram is shown with the two protomers of disulfide-linked VEGF shown in orange and purple, and Ig-like domain 2 of Flt1 shown in green. b) Structural model of the human growth hormone bound to the ligand binding domain of the corresponding receptor. The monomeric human growth hormone hGH (red) binds in a nonequivalent manner to two molecules (green and blue) of the ligand binding domain of the hGH receptor. The membrane bilayer is modeled below.

Formation of Heterodimers

An aspect of ligand-induced oligomerization of receptors of regulatory importance is the possibility of forming heterodimers. Protein families composed of closely related members can be identified for a number of growth factors and corresponding receptors.

Analogous to heterotypical dimerization of transcription factors (see Section 1.4.5), heterologous dimerization is observed within different members of a receptor family. A certain growth hormone can thus bind to and activate different dimeric combinations of the members of a receptor family. Figure 8.6 shows the possibilities for heterodimerization of receptors, using the PDGF receptor as an example.

Heterodimerization of receptor molecules is a mechanism that can increase and modulate the diversity and regulation of signal transduction pathways. Since the various members of a receptor family differ in the exact structure of the autophosphorylation sites and the other regulatory sequences, it is assumed that activity and regulation are different for the various combinations of receptor subtypes. Tissue-specific expression of receptor subtypes enables the organism to process growth hormone signals in a differential way.

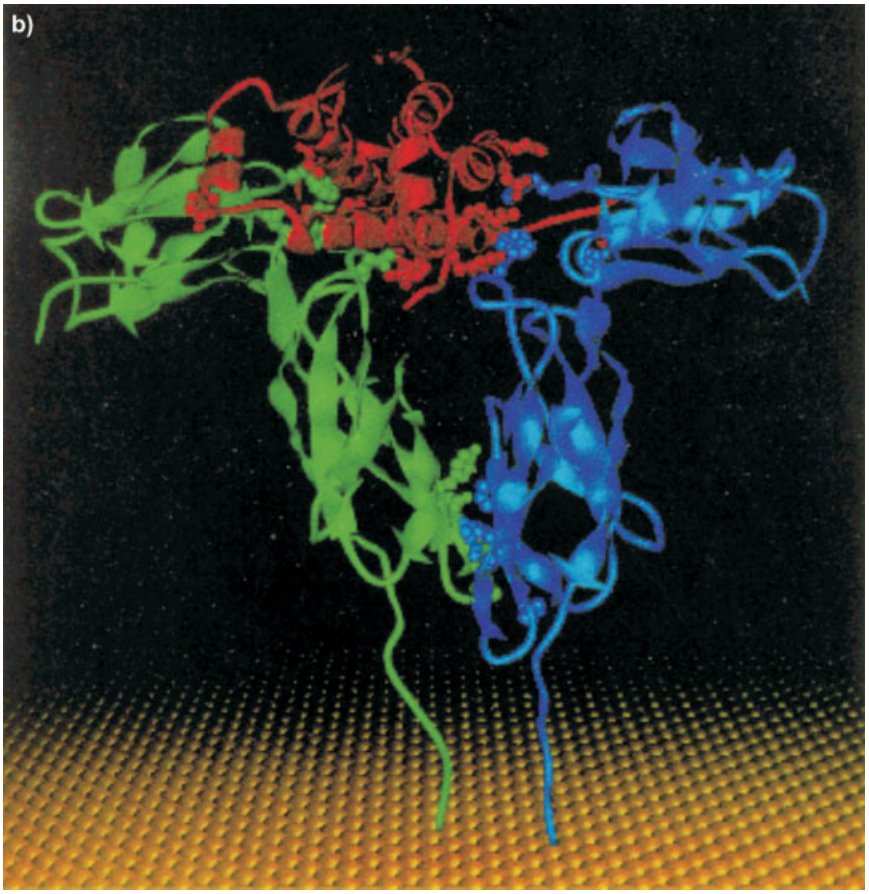


Fig. 8.5

8.1.3 Structure and Activation of the Tyrosine Kinase Domain

Typically, two processes are required for activation of RTKs: stimulation of intrinsic catalytic activity and creation of binding sites for downstream effector proteins. Ligand-induced autophosphorylation on tyrosine residues serves for both processes.

Enhancement of the catalytic activity is achieved by autophosphorylation in the activation segment within the kinase domain. It is generally assumed that this autophosphorylation takes place by a trans mechanism. Accordingly, two neighboring Tyr kinase domains in the receptor oligomer perform a mutual phosphorylation (see Fig. 8.3).

The docking of downstream signaling proteins is mediated by autophosphorylation on tyrosines located outside the kinase domain, namely in regions immediately following the transmembrane segment and in the C-terminal region. Autophosphoryla-

tion in these regions creates docking sites for modular domains of downstream signaling proteins that recognize phosphotyrosine residues in specific sequence contexts (see Section 8.2).

By binding to the phosphotyrosine docking sites, the effector proteins next in sequence are recruited into the signaling process. The effector proteins may carry enzyme activity themselves and be activated by tyrosine phosphorylation. They may also function as adaptor molecules, functioning to pass the signal on to other components of the signaling pathway. With the help of adaptor molecules, other signal proteins are directed to the activated receptor and to the cell membrane, and are thus incorporated into signal transduction.

What is the purpose of the docking of the effector proteins? In many cases, the effector proteins are substrates for Tyr-phosphorylation catalyzed by the tyrosine kinase activity of the receptor, and this phosphorylation is required for further signal transmission. The colocalization of the substrates and the kinase via the docking sites helps to increase the efficiency and specificity of phosphorylation, and it is often a first step for the assembly of larger signaling complexes in the vicinity of the activated receptor. This is mainly true for effector proteins with adaptor and scaffolding function that are often multivalent and can assemble several signaling proteins. *Autophosphorylation* and *phosphorylation of substrate proteins* are therefore essential elements of signal transduction via receptor tyrosine kinases.

Structure of the Tyr Kinase Domain of Receptor Tyrosine Kinases

The crystal structures of the cytoplasmic domains of several receptor tyrosine kinases have been reported, providing a structural basis for the explanation of the stimulation of kinase activity by activation segment phosphorylation (review: Hubbard and Till, 2000). On the example of the insulin receptor, for which most data are available, the principles of activation by autophosphorylation are presented in the following. The insulin receptor is a heterotetrameric receptor tyrosine kinase of an $\alpha_2\beta_2$ structure (see Fig. 8.2). The α subunit is completely extracellular and is bound to the β chain via disulfide bridges. The β chain has a transmembrane portion, and the tyrosine kinase domain is on the cytosolic side. On binding insulin on the extracellular side, the Tyr kinase activity of the β chain is activated, and autophosphorylation of a total of seven Tyr residues takes place in the *cytoplasmic domain*. Consequently, an effector protein, *insulin receptor substrate* (IRS, see Section 8.5), binds to phosphotyrosine residues of the activated receptor via its phosphotyrosine-binding domain. IRS now becomes phosphorylated by the insulin receptor at tyrosine residues that act as docking sites for the SH2 domains of other assigned proteins. Targeted mutations in the region of the catalytic center of the insulin receptor have shown that the Tyr kinase activity is an essential function in signal transduction via insulin.

A total of seven tyrosine residues become phosphorylated during autophosphorylation. Two of these are located in the vicinity of the transmembrane element, three are in the activation segment of the catalytic domain, and a further two in the region of the C terminus.

The crystal structure of a 306 amino acid fragment of the β chain of the insulin receptor comprising the catalytic domain indicates that the Tyr kinase domain has

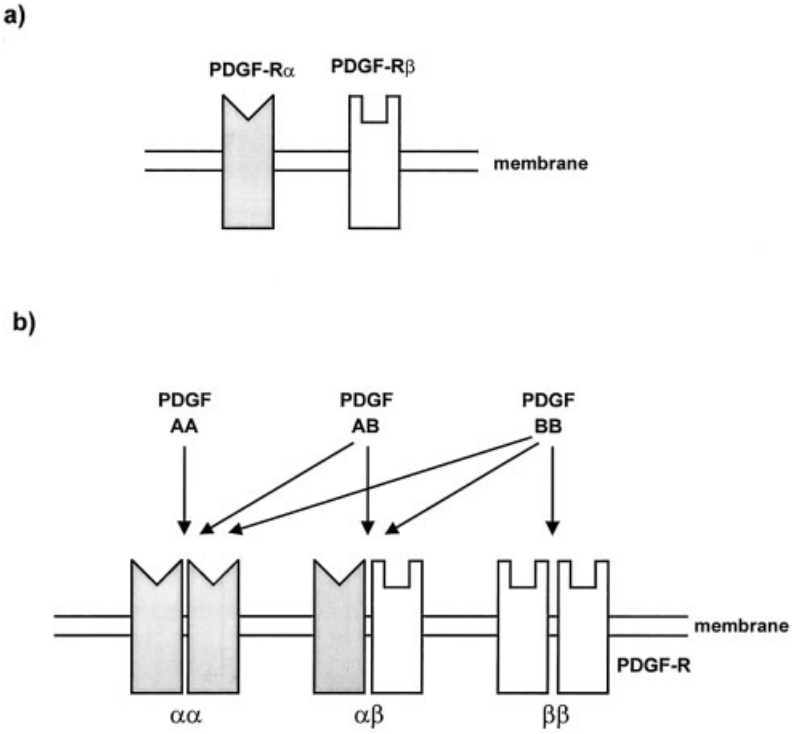


Fig. 8.6 Heterodimerization of PDGF receptors. a) There are α and β subtypes of platelet derived growth factor receptor (PDGF-R); these are induced by ligand binding to form homodimers and heterodimers. b) Platelet derived growth factor (PDGF) is a dimeric growth factor, composed of chains A and/or B. The protein may exist as a homodimer (AA, BB) or heterodimer (AB). The AA homodimer of PDGF binds to the $\alpha\alpha$ dimer of PDGF-R, AB binds to the $\alpha\alpha$ and $\alpha\beta$ types, BB binds to all three combinations.

a similar construction to that of the catalytic domain of the Ser/Thr-specific protein kinases (see Section 7.2). Like the latter, the Tyr kinase domain is composed of two lobes and contains most the structural features relevant for protein kinase activity and regulation (see Fig. 7.3). As for many other protein kinases, the activation segment is of great importance for regulation of the Tyr kinase activity by phosphorylation. Three tyrosine residues of the activation loop become phosphorylated during autophosphorylation, inducing a striking conformational change. In the unphosphorylated insulin receptor, Tyr 1162 is bound in the active site, obstructing access of the ATP and the protein substrate to the active site. Autophosphorylation of the activation segment brings about a dramatic repositioning of the segment (Fig. 8.7). The steric hindrance to substrate binding (Mg-ATP and peptide substrate) is removed, and the residues involved in substrate binding and catalysis are now properly positioned. The alternative conformation of the activation segment is fixed by ionic interactions involving the phosphotyrosine residues.

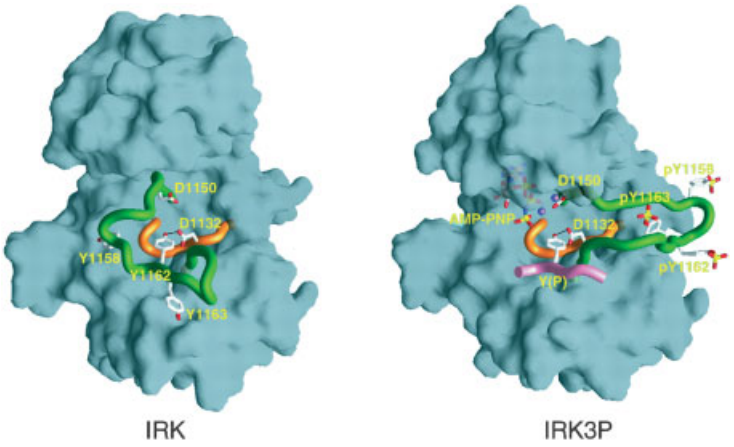


Fig. 8.7 Comparison of the activation loop conformation in unphosphorylated insulin receptor kinase (IRK) and tris-phosphorylated IRK (IRK3P). The figure illustrates the dramatic repositioning of the activation loop upon autophosphorylation. Note, that in IRK, Tyr-1162 is positioned in the active site competing with protein substrates. The activation loop is shown in green, the catalytic loop in orange, and the peptide substrate in pink. The rest of the protein in each case is represented by a semitransparent molecular surface. Also is shown the ATP analog AMP-PNP, which is partiall masked by the amino-terminal lobe of IRK3P. Carbon atoms are shown in white, nitrogen atoms in blue, ooxygen atoms in red, phosphorous atoms in yellow, and magnesium ions in purple. Hydrogen bonds between the substrate tyrosine Y(P) and Asp-1132 (IRK3P) are indicated by black lines. From Hubbard and Till, (2000).

From the temperature factors derived from the crystallographic analysis, it can be concluded that the activation segment is quite mobile and can exist in different conformations in solution. The equilibrium between the multiple conformations appears to be shifted to a catalytically active conformation in the phosphorylated state that is fixed by interactions with other structural parts of the catalytic center.

How does binding of insulin to the extracellular domain trigger phosphorylation of the activation segment? Structural information on this point is not yet available. It is postulated that an allosteric transition of the heterotetrameric receptor takes place on insulin binding, bringing about a change in the mutual configuration of the β chains on the cytosolic side. Consequently, the phosphorylation sites and the active centers of two β chains are orientated so that a mutual phosphorylation is possible. The inhibitory Tyr1162 is removed from the active center during this process, and the substrate-binding sites become accessible. The conformational change permits a *trans-phosphorylation of both β chains*, which includes Tyr phosphorylation in the activation segment. As a result, the equilibrium of the receptor conformations is shifted to the side of the active form. According to this theory, phosphorylation of the activation segment has a key function in receptor activation. All receptor tyrosine kinases identified so far have one or more tyrosine residues in the activation segment. With few exceptions (e. g., EGF receptor), receptor tyrosine kinases are phosphorylated on one to three

tyrosine residues in the activation segment, and these phosphorylation events are critical for receptor activation.

8.1.4 Effector Proteins of the Receptor Tyrosine Kinases

Autophosphorylation of receptor tyrosine kinases has a double effect: The tyrosine kinase activity undergoes autoactivation by phosphorylation of Tyr residues localized in the catalytic domain. In addition, Tyr residues that lie outside the active center are phosphorylated. The phosphotyrosine residues thereby created serve as *binding sites* for effector molecules next in the sequence of the signal transduction pathway (see Fig. 8.8a).

The phosphotyrosine residues of the activated receptors are attachment points for effector proteins that possess a phosphotyrosine-binding domain, such as the SH2 domain or the PTB domain (see Section 8.2). The effector protein is involved in the signal transduction pathway via interaction of its phosphotyrosine-binding domain with the phosphotyrosine residue (and neighboring sequence portions) of the activated receptor. Effector proteins may be enzymes or proteins that only possess adaptor function and link further proteins to the activated receptor via protein-protein interactions. The docking of signaling proteins to autophosphorylation sites provides a mechanism for assembly and recruitment of signaling complexes by receptor tyrosine kinases. Accordingly, activated receptor tyrosine kinases can be considered as a platform for the recognition and colocalization of a specific complement of signaling proteins.

Activated receptor tyrosine kinases transduce signals to a variety of central intracellular signaling pathways. A multitude of effector proteins are recruited into signaling by receptor tyrosine kinases. Important examples are

- the p85 subunit of phosphoinositide-3-kinase (PI3-kinase) (see Section 6.6.1)
- phospholipase C-γ (see Section 5.6.2)
- nonreceptor tyrosine kinases of the Src family (see Section 8.3)
- p120 GAP, a GTPase activating enzyme of Ras signal transduction (see Chapter 9)
- the adaptor proteins Grb2 and Shc of Ras signal transduction and the scaffolding protein insulin receptor substrate, IRS1 (see Section 8.5 and Chapter 9).
- the Tyr-specific protein phosphatase SH-PTP2 (see Section 8.4).

These examples illustrate the diversity of receptor tyrosine kinase-initiated signal transduction pathways. Many of the signaling pathways activated by receptor tyrosine kinases ultimately lead to the activation of transcription factors, influencing central differentiation and developmental programs of the cell. Figure 8.8b summarizes the flow of signals from activated receptor tyrosine kinases through central signaling pathways.

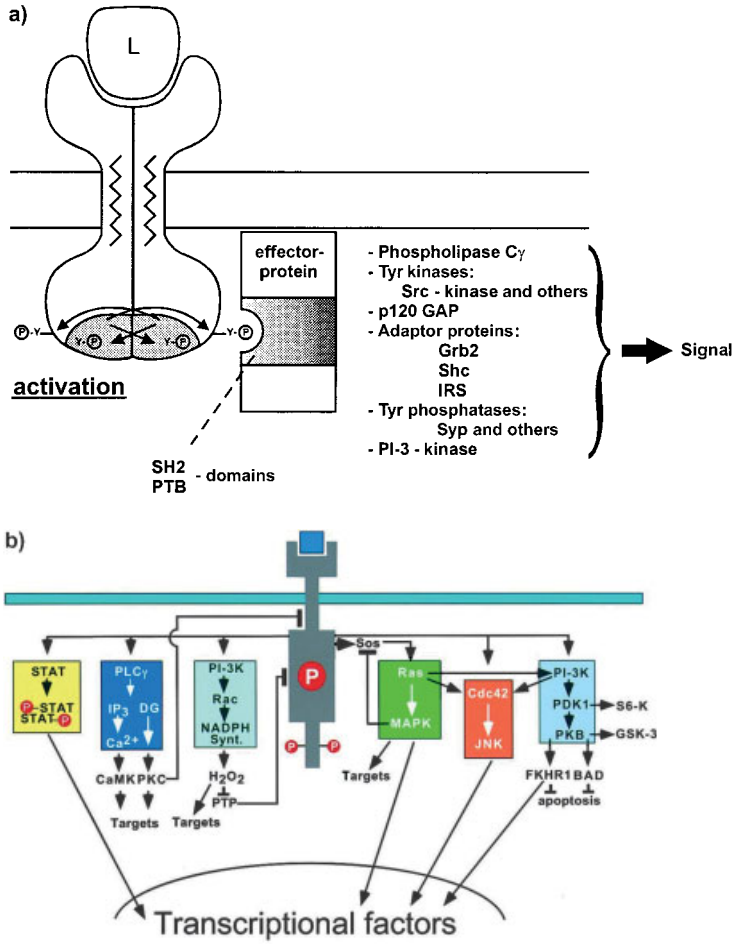


Fig. 8.8 A) Functions of autophosphorylation of receptor tyrosine kinases. Autophosphorylation of receptor tyrosine kinases takes place in trans, i. e., between neighboring protomers of the receptor. The catalytic domain of the receptor is shown as a shaded segment. As a consequence of autophosphorylation, the intrinsic tyrosine kinase activity of the receptor is stimulated. Effector proteins can also bind to the activated receptor. Binding takes place with specific phosphotyrosine binding domains (SH2 or PTB domains) at phosphotyrosine residues of the activated receptor. A critical factor for further signal transduction is the membrane association of the effector proteins that enter into binding with the activated receptor. Details of the effector proteins can be found as follows: phospholipase Cγ: 5.6.2; Src kinase: 8.3.2; p120 GAP: 9.4; Grb2, Shc, IRS: 8.5; PI3-kinase: 6.6.1; Syp tyrosine phosphatase: 8.4

B) Signaling pathways activated by receptor tyrosine kinases. Different signaling pathways are presented as distinct signaling cassettes (coloured boxes). Not all known components of a given pathway are included. Examples of of stimulatory and inhibitory signals for the different pathways are also shown. The signaling cassettes presented in the figure regulate the activity of multiple cytoplasmic targets. However, the Ras/MAPK, STAT, JNK and PI3-kinase signaling pathwaysalso regulate the activity of transcriptional factors by phosphorylation and other mechanisms.

STAT: see Section 12.2.2; PLC: see Chapter 6; PI3-K see Section 6.6.1; Ras: see Chapter 9; MAPK, Cdc42/JNK: see Chapter 10; PTP: protein tyrosine phosphatase; FKHR: Forkhead transcription factor; S6-K ribosomal protein S6 kinase; GSK-3: gylcogen synthase kinase 3. After Schlessinger (2000).

PI3-Kinase Pathway

A subgroup of PI3-kinase enzymes containing a p85 subunit can bind via its SH2 domain to the autophosphorylated receptor or to phosphorylated docking proteins assembled at the activated receptor such as IRS1. Activated PI3-kinase generates PtdIns(3,4,5)-P, which mediates the membrane association of a variety of signaling proteins (see Section 6.6). One important response is stimulation of cell survival. Furthermore, many of the biological effects of insulin are mediated via activation of the PI3-kinase pathway.

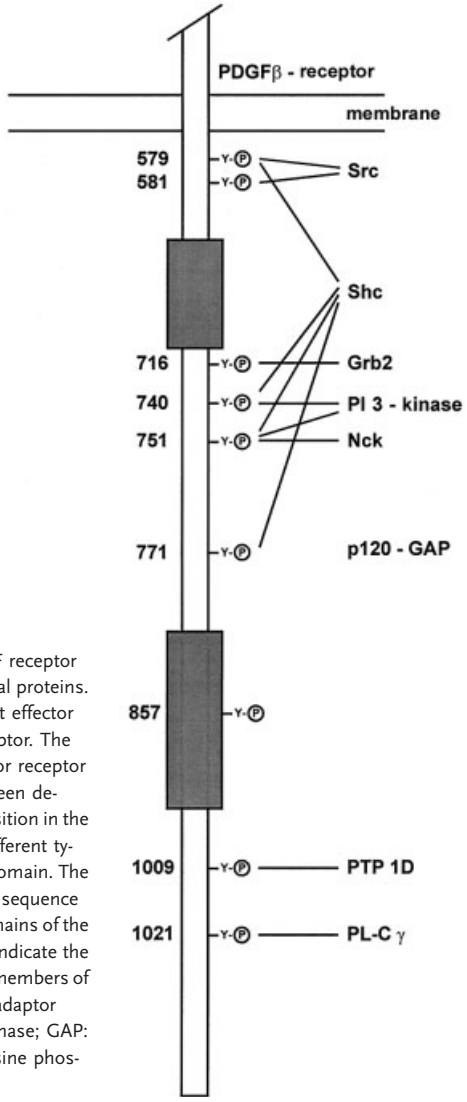


Fig. 8.9 Phosphotyrosine residues in the PDGF receptor and specificity of binding of SH2-containing signal proteins. The figure illustrates the diversity of the different effector proteins that can interact with an activated receptor. The tyrosine residues of platelet derived growth factor receptor (PDGF-R), for which autophosphorylation has been demonstrated, are designated according to their position in the receptor sequence. PDGF-R has at least nine different tyrosine phosphorylation sites in the cytoplasmic domain. The phosphotyrosine residues are found in different sequence environments and are recognized by the SH2 domains of the assigned effector proteins. The filled rectangles indicate the two-part tyrosine kinase domain of PDGF-R. Src: members of the Src tyrosine kinase family; Sh2, Grb2, Nck: adaptor proteins; PI(3) kinase: phosphatidyl inositol-3 kinase; GAP: GTPase activating protein; PTP-1D: protein tyrosine phosphatase 1D; PL-Cγ: phospholipase C-γ.

Ras Pathway (see Chapter 9)
One way of activating Ras signaling requires the recruitment of adaptor proteins like Grb2 and Shc to the activated receptor (see Section 9.5). A multitude of signals can be delivered via Ras proteins, resulting in activation of the MAK kinase cascade and the activation of transcription factors.

Phospholipase Cγ (see Sections 5.6.2 and 6.4)
Activation of receptor tyrosine kinases leads to stimulation of phospholipid metabolism and to the generation of multiple second messengers. Phospholipase Cγ binds through its SH2 domain to phosphotyrosine sites on the receptor molecules and is thereby activated. As a consequence of phospholipase Cγ activation, Ca^{2+} and diacylglycerol signals are produced leading, e. g., to the activation of protein kinase C isoenzymes and of CaM kinases.

Already at the level of a given receptor, significant *branching* and *variability* of signal transduction is possible. Which effector protein is bound to an activated receptor tyrosine kinase depends on the nature of the SH2 domain of the effector protein and on the sequence environment of the phosphotyrosine residue. The receptor tyrosine kinase typically has several autophosphorylation sites with different neighboring sequences, so that every phosphotyrosine residue of the receptor tyrosine kinase may serve as the binding site for a different effector molecule. Figure 8.9 illustrates the diversity of effector proteins that can interact with a receptor type, using the PDGF receptor as an example.

The specific interaction between the phosphotyrosine residues of the activated receptor and the SH2 domain of the effector protein is the basis for the activation of effector proteins for further signal transduction. Like the nature of the effector proteins, the mechanism of their activation is also very variable. Important mechanisms are (review: Schlessinger, 2000)

- Phosphorylation of the effector molecule at Tyr residues
  Example: phospholipase C-γ
- Induction of a conformational change in the effector molecule
  Example: PI3-kinase
- Translocation of the effector molecule to the plasma membrane
  Example: Grb2-Sos, Shc-Grb2 (see Section 9.5).

8.1.5 Attenuation and Termination of RTK Signaling

Signaling by RTKS must be tightly regulated and properly balanced in order to produce the normal cellular responses (review: Pawson, 2002; Schlessinger, 2000). We know of many examples where aberrant expression or dysfunction of RTKs is responsible for developmental disorders and diseases including tumor formation. The cell uses several mechanisms for the attenuation and termination of RTK signaling induced by stimulatory ligands (Fig. 8.10).

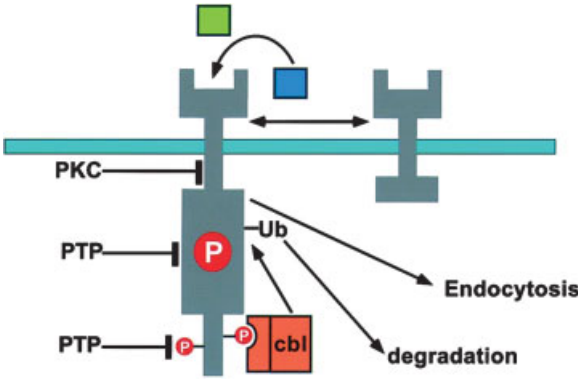


Fig. 8.10 Mechanisms for attenuation and termination of signaling by receptor tyrosine kinases, RTKs. In several cases the activity of RTKs can be negatively regulated by ligand antagonists or by hetero-oligomerization with naturally occurring dominant interfering receptor variants. The PTK activity of EGFR is attenuated by protein kinase C (PKC)-induced phosphorylation at the juxtamembrane region. Dephosphorylation of key regulatory pTyr residues by protein tyrosine phosphatases may inhibit kinase activity or eliminate docking sites. An important mechanism for signal termination is via receptor endocytosis and degradation mediated by ubiquitin ligases as e. g. the Cbl protein (see also Fig. 2.14).

Antagonistic Ligands, Hetero-dimerization

In several systems, natural antagonistic ligands have been identified that bind on the extracellular side to the receptor and inhibit receptor activation. Furthermore, certain tissues express naturally occurring receptor variants that are deficient in tyrosine kinase activity. Expression of these mutants may lead to dominant negative inhibition of full-length receptors through formation of inactive hetero-dimers or hetero-oligomers.

Inhibition of RTK Activity

Protein kinase C-mediated phosphorylation of the EGF receptor on Ser/Thr residues located on the cytoplasmic domain results in inhibition of its tyrosine kinase activity and in inhibition of EGF binding to the extracellular domain. This phosphorylation appears to provide a negative feedback mechanism for control of receptor activity. In another negative control mechanism, inhibitory proteins like SOCS (suppressor of cytokine signaling) bind via their SH2 domains to phosphotyrosine residues in the tyrosine kinase domain of the receptor and directly inhibit the RTK activity.

Inhibition by Protein Tyrosine Phosphatases

The activity of RTKs is continuously subject to control by PTPs (see Section 8.4) that can dephosphorylate and thereby deactivate RTKs that have been autophosphorylated because of binding of a stimulatory ligand.

Endocytosis and Degradation

Binding of extracellular ligands to RTKs often results in rapid endocytosis and degradation of both the receptor and the ligand, attenuating the signal generated at the cell surface. The oncogenic adaptor protein Cbl has been shown to play a role in the degradation of the EGF receptor and PDGF receptor. Cbl contains a tyrosine kinase-binding domain mediating binding of the activated receptor and RING finger domain that functions as a ubiquitin ligase, mediating ubiquitination and proteasomal degradation of the receptor (see Fig. 2.14).

8.2 Protein Modules as Coupling Elements of Signal Proteins

Starting from an activated receptor tyrosine kinase, further conduction of the signal takes place with the help of specific protein-protein interactions between the activated receptor and one or more effector proteins next in the sequence. In many cases, the effector molecules pass the signal on to other proteins of the signaling pathway, forming chains of signal proteins in sequence. Specific protein-protein interactions are the basis of the coupling of signal proteins (review: Pawson, 2002). These function mainly to bring about close spatial configuration of signal-carrying enzymes with their substrates, by leading a substrate protein to the catalytic center or by targeting an enzyme usually located in the cytosol to the cell membrane, where it has direct access to its substrates.

The cell uses defined structural elements for communication between different proteins of a signal transduction pathway and for targeting signaling proteins to membranes; these are found in the form of self-folding protein domains in many signal-transmitting proteins. The protein domains are modules of 60 – 100 amino acids that – with a common basic structure – occur in very different proteins in slightly modified forms. The protein modules mediate protein-protein interactions or protein-membrane interactions in signaling pathways (Fig.8.11). They are used to associate proteins of a signaling pathway into larger signaling complexes or to target them to the vicinity of the cell membrane (reviews: Pawson and Scott, 1997; Sudol, 1998).

Two points are of particular importance for the coupling function of the protein modules. The first is the variability of the protein modules. For a particular basic motif of a module, there are generally many *variants* that have slightly different binding specificities and are thus assigned to different structural motifs in the target protein. This results in great variability and diversity of coupling. Another functionally important aspect of coupling of signal proteins is the occurrence of *several* protein modules in a protein (Fig. 8.11). If a signal protein has multiple binding valence for different effector proteins, networks of interacting proteins can be created that contribute greatly to specificity and diversity of signal transduction and permit linking of different signaling pathways.

At present, a number of structural motifs have been described for signal proteins to which specific coupling functions in signal transduction are attributed. The functions and targets of the most important signaling modules are summarized in Fig. 8.12.

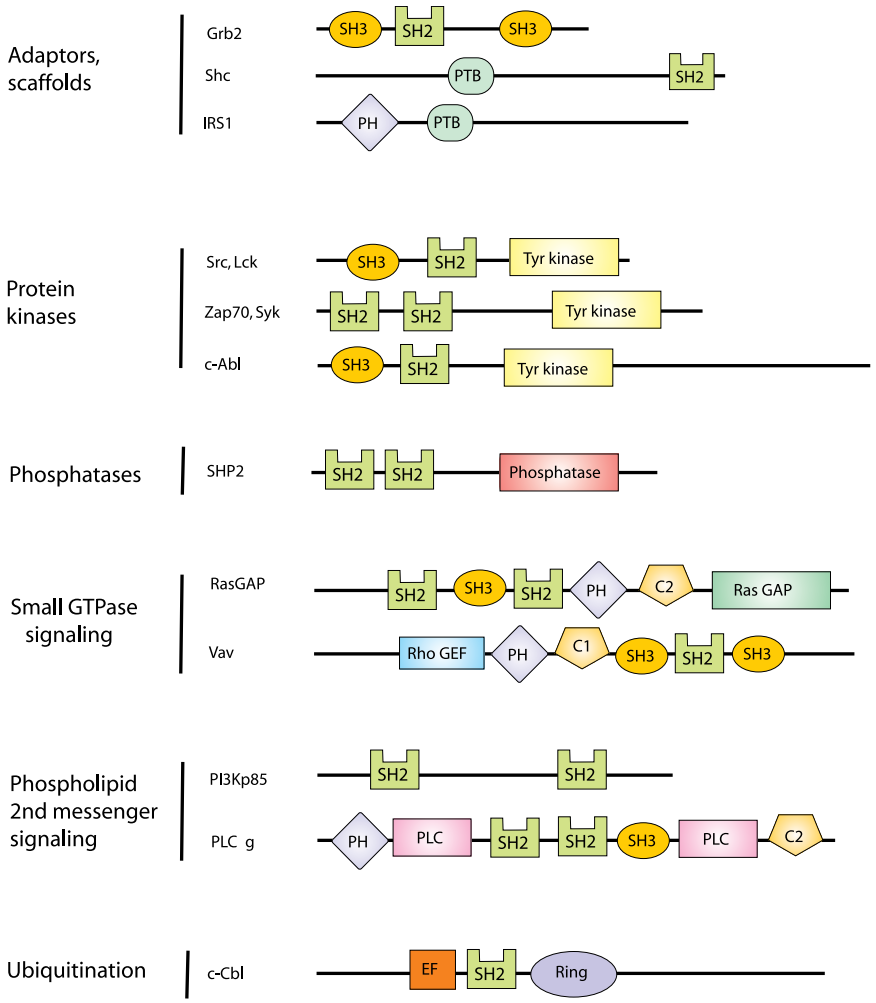


Fig. 8.11 The modular nature of signaling proteins.Representative members from various SH2 domain families and the positional organization of these domains are illustrated. For explanation of protein modules see text. C1, C2, cysteine-rich domains; PI3Kp85: p85 subunit of PI3K.

8.2.1 SH2 Domains

The SH2 domains were first discovered as a sequence motif showing homology with a sequence of the Src tyrosine kinase: hence the name SH, from Src homolog. Src tyrosine kinase was the first known tyrosine kinase. The enzyme was discovered in avian retroviruses that induce sarcomas. Later it became apparent that Src tyrosine kinase also belongs to the normal enzyme repertoire of mammalian cells and is involved in the regulation of cell growth and differentiation (see Section 8.3).

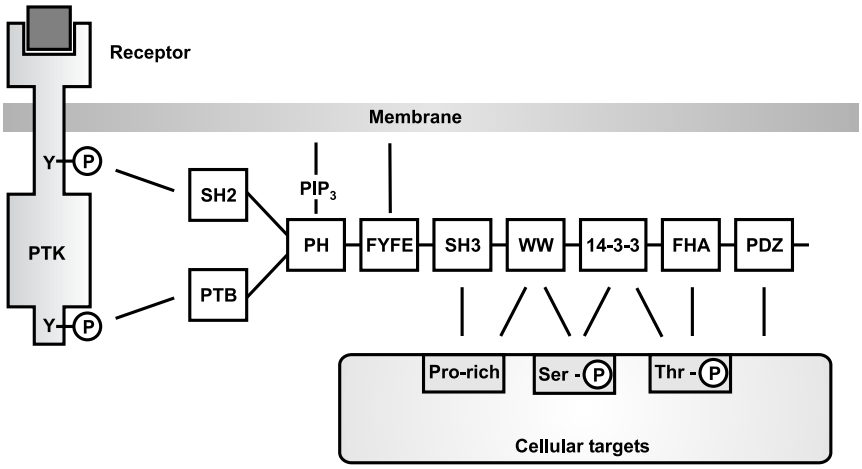


Fig. 8.12 Targets of important signaling modules.

Binding Specificity and Structure of SH2 Domains

The SH2 domain functions to specifically recognize phosphotyrosine residues in target proteins and to bind these (review: Pawson et al., 2001). The presence of a Tyr phosphate grouping is obligatory for binding of an SH2 domain to a target protein or to a model peptide. In addition, the neighboring sequence is crucial. The sequence environment of the phosphotyrosine residue defines the binding substrate of a particular SH2 domain and differentiates the binding preference of different SH2 domains. Specificity is determined by sequences of 1 – 6 aminoacids located C-terminal to the phosphotyrosine residue. In contrast, the other known phosphotyrosine-binding motif, the PTB domain, recognizes sequences N-terminal to the phosphotyrosine (see below). The great variability of SH2 domains and their substrates is emphasized by the observation that two SH2 domains occur in many signal proteins. These mostly have different substrate-binding preferences.

SH2 domains can be divided into at least five classes (1A, 1B, 2, 3, 4), differing in the sequence requirements of the substrate (Cantley and Songyang, 1997). As an example, Fig. 8.13 schematically illustrates interactions in the complex between the SH2 domain of the Src kinase, which belongs to SH2 group 1A, and a Tyr-phosphorylated peptide.

Crystallographic analysis of the complex shows that the phosphate residue is bound in a deep pocket of the SH2 domain, at the end of which an invariant Arg residue is located which contacts the negatively charged phosphate by a two-pronged interaction. It can be estimated that a phosphoserine or phosphothreonine residue would be too short to enter into a similar interaction with the Arg residue. The neighboring sequences of the phosphotyrosine residue are decisive for binding specificity of an SH2 domain. At this point, the structure shows that particularly the isoleucine residue at position +3 relative to phosphotyrosine is bound in a very specific manner in a pocket of the SH2 domain (Fig. 8.13). Binding of the peptide to the SH2 domain in Src kinase has therefore been compared to binding of a two-pole plug in a com-

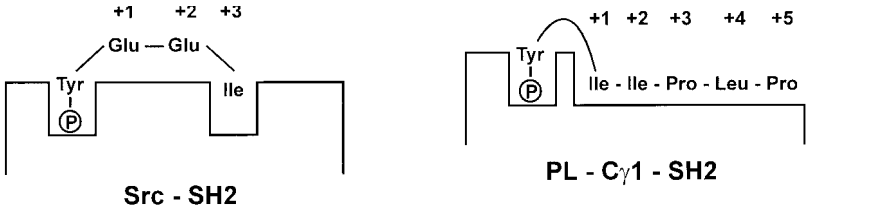


Fig. 8.13 Recognition of phosphotyrosine-containing substrate peptides by SH2 domains of Src kinase and phospholipase Cγ1. Binding of phosphotyrosine-containing peptides to SH2 is shown schematically, based on crystal structures of the complexes. The SH2 domain of Src kinase has a basic binding pocket for the phosphotyrosine residue and a hydrophobic pocket for the isoleucine residues at position +3 of the peptide substrate. The SH2 domain of PL-Cγ1 has a hydrophobic binding surface to which the C-terminal part of the peptide P-Tyr-Ile-Ile-Pro-Leu-Pro-Asp binds.

plementary socket, where one of the poles is phosphotyrosine and the other is the amino acid at position +3.

Function of the SH2 Domain

The interaction between the SH2 domain of a signal protein and a phosphotyrosine-containing protein ligand serves to pass the signal on in a signal chain. It is the complementarity between a given SH2 domain and the sequence context of a phosphotyrosine residue in the target protein that determines the specificity of signal transmission at this point. Often multiple phosphotyrosine residues with different sequence contexts are found in signaling proteins. An example is the β-subtype of the PDGF receptor on which several tyrosine phosphorylation sites have been identified allowing the docking of at least eight different effector molecules. The same receptor can be involved in very different signaling pathways, as shown by this example. Which pathway is used will depend on the availability and activity of the different effector proteins, a situation regulated in a cell-specific and tissue-specific manner.

Phosphotyrosine residues as targets of SH2 domains are found in a variety of signaling proteins. As well as the receptor tyrosine kinases and the cytokine receptors, the T-cell receptors and the nonreceptor tyrosine kinases utilize phosphotyrosine-SH2 interactions for signal transmission. Other examples include adaptor or docking proteins like Grb2 and Shc, phospholipase Cγ, and transcription factors like the Stat proteins (see Section 11.2.2).

Activation of signal proteins by phosphotyrosine-SH2 interactions can be achieved in different ways, which are discussed below.

- **Activation by membrane localization**
  Via binding of an SH2-containing signal protein to an activated receptor tyrosine kinase, the signal protein is brought to the membrane and into the vicinity of the corresponding target protein or substrate. Examples are phospholipase Cγ and PI3-kinase, which have substrates in the phospholipid membrane. The same is true for the GTPase-activating protein of the Ras pathway and for the Grb2-Sos complex,

both of which have the Ras protein as target protein (see Chapter 9). In these cases, the target proteins are localized in or at the membrane, and activation of the signaling pathway initiates translocation of the cytosolic signal protein to the membrane, where it is in the immediate vicinity of the substrate or target protein.

- **Activation by tyrosine phosphorylation**
  Many SH2-containing signal proteins are brought, via interaction of their SH2 group with phosphotyrosine residues, into the neighborhood of the catalytic center of the tyrosine kinase and are themselves substrates for tyrosine phosphorylation. By this mechanism, new attachment sites can be generated for other SH2-containing proteins within SH2-containing signal proteins. In this way, several components of a signaling pathway can be sequentially linked.
- **Activation by a conformational change**
  Several cases are described in which binding of an SH2-containing enzyme to an activated receptor tyrosine kinase leads to increased catalytic activity of the enzyme. One example is the PI3-kinase. Binding of the regulatory subunit p85 to a tyrosine-phosphorylated PDGF receptor causes conformational changes in p85 that are transmitted to the catalytic p110 subunit and stimulate the PI3-kinase activity. Src kinase is another example of a protein kinase that is allosterically controlled by SH2-phosphotyrosine interactions (see Section 8.3.2).

8.2.2
Phosphotyrosine-binding Domain (PTB Domain)

The phosphotyrosine-binding (PTB) domains harbor – like the SH2 domains – specific binding sites for phosphotyrosine-containing peptides. PTB domains are found particularly in proteins that have a docking or adaptor function by recruiting additional signaling proteins to the vicinity of an activated receptor. Examples are the adaptor molecule Shc (see Section 8.5), which has an SH2 domain of class 3 in addition to the PTB domain, and the insulin receptor substrate IRS1. PTB domains recognize phosphotyrosine residues in context with sequence sections toward the N terminus. However, there is a great heterogeneity in the binding properties of different PTB domains. Certain PTB domains bind to nonphosphorylated peptide sequences, while others recognize both phosphorylated and nonphosphorylated sequences (reviews: Margolis, 1999, Yan et al., 2002).

8.2.3
SH3 Domains

SH3 domains occur in signal proteins that are involved in Tyr kinase signaling pathways (review: Macias et al., 2002). They are also found in proteins of the cytoskeleton and in a subunit of the neutrophilic cytochrome oxidase. Ligand binding at SH3 domains takes place via *Pro-rich sequences* of ca. 10 amino acids. The sequence X-P-p-X-P is a consensus sequence for SH3 ligands, in which the two proline residues P are invariant; X is usually an aliphatic residue and p is often a Pro residue. The structural

determination of SH3 domains with bound Pro-rich peptides has shown that the Pro-rich section of the ligand is bound as a left-handed polyproline type II helix with three amino acid residues per turn. The polyproline type II helix was described for polyproline some time ago.

Like the SH2 domains, there are many different SH3 domains. The different SH3 domains demonstrate differing binding preferences for Pro-rich sequences, the specificity being determined by the neighboring residues of the invariant proline.

A general function of the SH3 domains is the binding of Pro-rich sequences in target proteins of a signaling pathway. SH3 domains are thus coupling modules in signaling pathways. The biological importance of the SH3 domains is emphasized by the observation that deletion of the SH3 domains of the cytoplasmic tyrosine kinases Abl and Src leads to a significant increase in the tumor-transforming potential of both tyrosine kinases. The following principal functions can be attributed to SH3 domains:

- **Mediation of specific subcellular localization**
  SH3 domains are found in many proteins associated with the cytoskeleton or with the plasma membrane. Examples are the actin-binding protein α-spectrin and myosin Ib.
- **Regulation of enzyme activity**
  An example of regulation of enzyme activity via SH3 domains is the negative regulation of the activity of Src tyrosine kinase by SH3-mediated interactions (see Section 8.3).
- **Contribution to substrate selection of tyrosine kinases**
  Tyrosine kinases demonstrate – as do many Ser/Thr-specific protein kinases – intrinsically low substrate specificity, and it is therefore difficult to formulate consensus sequences for phosphorylation sites of the substrates of tyrosine kinases (Songyang and Cantley, 1995). One means of increasing the selectivity of Tyr phosphorylation seems to be to use SH3 domains for specific coupling of substrates to Tyr kinases (Fig. 8.14). The adaptor proteins Crk, Grb2 and Nck are specifically phosphorylated by the Abl tyrosine kinase at Tyr residues. All three adaptor proteins possess SH3 domains that can bind to Pro-rich sequences of the Abl tyrosine kinase. This interaction mediates tight binding of the substrate to the tyrosine kinase and enables an effective tyrosine phosphorylation. Other possible substrates that have no SH3 domain are excluded from Tyr phosphorylation since they cannot enter into tight binding with the tyrosine kinase.
  The specific function of the SH3 domain is based on *increased substrate specificity* of tyrosine kinases in this case. In classical enzymes, the substrate-binding site and the catalytic center are close together, and the substrate-binding site is generally highly specific for a particular substrate. The situation is different for tyrosine kinases. Here, the substrate-binding site near the catalytic center shows moderate selectivity. The specificity of the reaction is increased, however, by mediation of association with the tyrosine kinase via additional structural elements of the substrate, namely its SH3 domain. This example highlights an important function of signaling modules. The interaction between a signaling module of the tyrosine

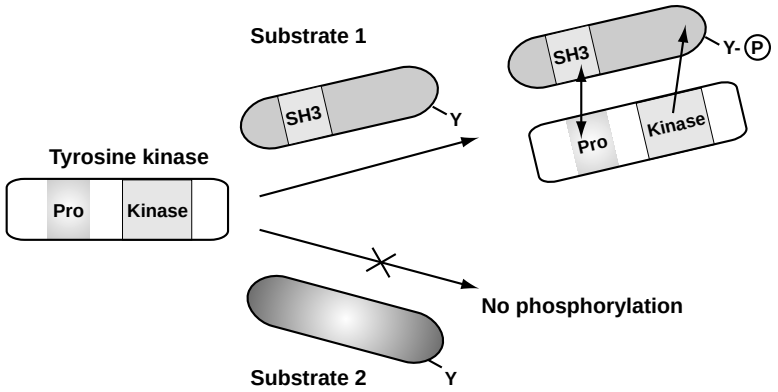


Fig. 8.14 SH3 domains and specificity of tyrosine kinases. Binding of the SH3 domain of the substrate protein S1 to the Pro-rich sequence of the tyrosine kinase increases affinity of the tyrosine kinase for its substrate and favors its phosphorylation. Proteins (S2) that possess Tyr phosphorylation sites but do not have SH3 domains are not converted by the tyrosine kinase, or only to a much lower extent.

kinase with a target motif on the substrate increases the specificity of the reaction and contributes to substrate recognition, bringing high selectivity to an otherwise rather unspecific catalytic reaction.

8.2.4 Membrane-targeting Domains: Pleckstrin Homology (PH) Domains and FYVE Domains

PH Domains

Pleckstrin homology (PH) domains comprise a large family of more than 200 domains and are found in many signal molecules such as Ser/Thr-specific protein kinases, tyrosine kinases, isoforms of phospholipase C (PL-Cβ, γ and δ), G nucleotide exchange factors, adaptor proteins, and proteins of the cytoskeleton (see also Fig. 8.11). Originally, the PH domain was discovered in the 47 kDa pleckstrin protein, which is the main substrate of protein kinase C in platelets.

PH domains bind phosphatide inositol derivatives and, because of this property, are able to mediate membrane association of signal proteins. A general membrane-anchoring function is therefore assigned to the PH domains (review: Lemmon and Ferguson, 2000) with phospholipids of the membrane serving as binding substrates (see Fig. 6.11). Signal-induced availability of phosphatidyl inositol lipids such as PtdInsP_3 thus permits regulated membrane-anchoring of PH-containing signal proteins (see Section 6.6).

There is great variation in the binding specificity of PH domains. While certain subtypes of PH domains bind specifically to PtdIns(4,5)P_2, another subset of PH domains binds preferentially to the products of the PI3-kinase reaction. However, most of the PH domains bind only nonspecifically and with weak affinity to phosphoinosi-

tides, and the nature of their natural substrates is open. Accordingly, it is also speculated that PH domains are used to mediate protein-protein interactions.

FYVE Domains

The FYVE domain is a cysteine-rich domain which binds two zinc ions and has binding specificity for PtdIns(3)P, one of the membrane-localized products of the PI3-kinase reaction (see Chapter 6). Proteins containing FYVE domains have been implicated, e. g., in membrane trafficking and in signaling by Smad proteins (review: Gillooly et al., 2001). It is assumed that binding of FYVE domain-containing proteins to PtdIns(3)P mediates their membrane association.

8.2.5 Phosphoserine/Threonine-binding Domains

Phosphorylation of proteins on Ser/Thr residues is one of the most common regulatory modifications of signaling proteins (see Chapters 2 and 7). Only recently has it been recognized that serine/threonine phosphorylation results in the formation of multiprotein signaling complexes through specific interactions between phosphorylated sequence motifs and the following phosphoserine/threonine-binding domains (review: Yaffe and Elia, 2001).

14-3-3 Proteins

14-3-3 proteins (review: Ferl et al., 2002) are a family of regulatory proteins which recognize phosphoserine/threonine residues in specific sequence contexts. These proteins are involved in the control of critical cellular functions such as cell cycle control, apoptosis, gene transcription, DNA replication and chromatin remodeling. Substrates of 14-3-3 proteins include, e. g., the Cdc25 phosphatase (see Section 13.6), the Raf kinase (see Section 9.6), the pro-apoptotic protein BAD (see Section 15.4), and histone deacetylase enzymes.

WW Domains

WW domains are signaling modules of ca. 40 amino acids that bind short proline-rich sequences such as PPLP or PPR motifs (review: Macias et al., 2002). A subset of WW domains, found, e. g., in the proline cis-trans isomerase Pin1, however, specifically binds to phosphoserine-Pro and phosphothreonine-Pro motifs. The Pin 1 protein has an essential role in mitosis. It is thought that Pin1 binding to phosphorylated mitotic proteins facilitates proline cis/trans isomerizations and subsequent conformational changes. For the Pin1 substrate Cdc25 phosphatase it has been shown that proline isomerization facilitates the subsequent dephosphorylation of phosphorylated Cdc25 protein by the protein phosphatase PP2A.

Forkhead-associated Domains

Forkhead-associated domains (FHA domains) were originally identified as conserved sequence elements within a subset of forkhead transcription factors and were subsequently found, e. g., in other transcription factors, in protein kinases, protein phosphatases, and kinesin motors (review: Durocher and Jackson, 2002). FHA domains comprise up to 140 amino acids, exhibit binding specificity toward phosphothreonine residues, and efficiently discriminate against phosphoserine residues.

8.2.6 PDZ Domains

PDZ domains were first identified in proteins of postsynaptic cells, and their designation comes from their occurrence in the proteins PSD-95, DlgA and ZO-1. In the meantime, PDZ domains have been found in many other proteins, particularly in proteins that form structures in the cell membrane (e. g., in ion channels) and in signal proteins (review: Hung and Sheng, 2002). PDZ domains recognize short peptide sequences with a C-terminal hydrophobic residue and a free carboxyl group, such as the E(S/T)DV motif at the C terminus of certain subunits of ion channels.

An important function of the PDZ domains lies in the formation of macromolecular associates at the cell membrane (review: Pawson and Scott, 1997). PDZ proteins can also provide a framework for clustering of ion channels within specific structures at postsynaptic membranes, known as postsynaptic density, PSD. Major organizers of the PSD appear to be PDZ-containing proteins, e. g., PSD-95, with distinct specificity for binding of downstream signaling proteins.

Many proteins contain multiple PDZ domains with various sequences that may show different binding specificities. In this way, a protein with multiple PDZ domains can help to organize different proteins in supramolecular complexes. An example is the InaD protein of *Drosophila*, which is composed exclusively of 5 PDZ domains with different binding specificities and to which different target proteins are assigned (Tsunoda et al., 1997, see Fig. 8.15). The target proteins are three proteins involved in the processing of light signals in the eye of *Drosophila*. During phototransduction, InaD associates via its distinct PDZ domains with phospholipase C-β, which is the target of rhodopsin-activated Gqα, with the calcium channel TRP, and with protein kinase C. The signaling complex formed allows efficient activation of the TRP channel by phospholipase C in response to stimulation of rhodopsin. Activated phospholipase C produces a PtdInsP_3 and a diacylglycerol signal and thereby induces opening of the TRP channel by a mechanism to be characterized. Furthermore, the presence of protein kinase C in the signaling complex provides for efficient deactivation of the TRP channel by phosphorylation. It is assumed that the InaD protein functions as an adaptor or scaffolding protein, which organizes light-induced signaling events into supramolecular complexes.

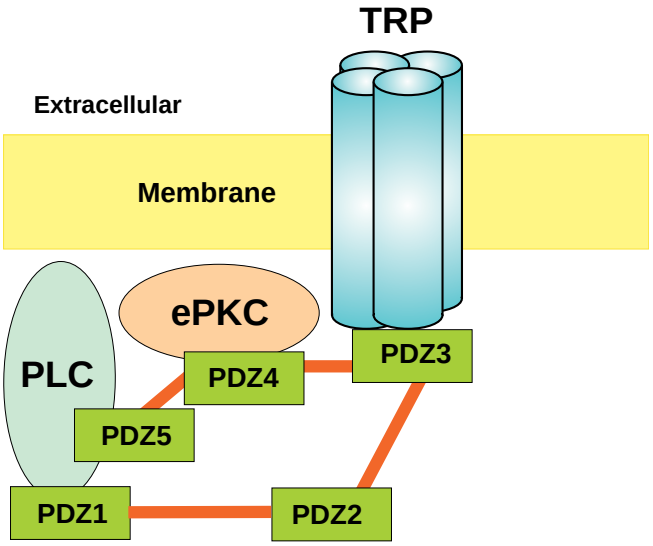


Fig. 8.15 A model of the organization of the InaD signaling complex. InaD is composed of 5 PDZ domains (PDZ1-5) which interact specifically with signaling proteins implicated in the vision process in *Drosophila*. TRP is a K^+-channel. PLC is a ß-type phospolipase C, which is the target of rhodopsin-activated Gq*α*. ePKC is an eye-specific protein kinase C which inactivates the Trp channel by phosphorylation (after Huber, 2001).

8.3 Nonreceptor Tyrosine-specific Protein Kinases

In addition to receptor tyrosine kinases, the cell also contains a number of tyrosine-specific protein kinases that are not an integral component of transmembrane receptors. These "nonreceptor" tyrosine kinases are localized in the cytoplasm at least occasionally or they are associated with transmembrane receptors on the cytoplasmic side of the cell membrane. They are therefore also known as *cytoplasmic tyrosine kinases*. The nonreceptor tyrosine kinases perform essential functions in signal transduction via cytokine receptors (see Chapter 11) and T cell receptors, and in other signaling pathways.

8.3.1 Structure and General Function of Nonreceptor Tyrosine Kinases

Permanent or transient association with subcellular structures, and variable subcellular distribution, are characteristic of the nonreceptor tyrosine kinases. These are intracellular effector molecules that can associate with specific substrates during the process of signal transduction and activate these by tyrosine phosphorylation to pass on the signal. Many of the functions of the nonreceptor tyrosine kinases are

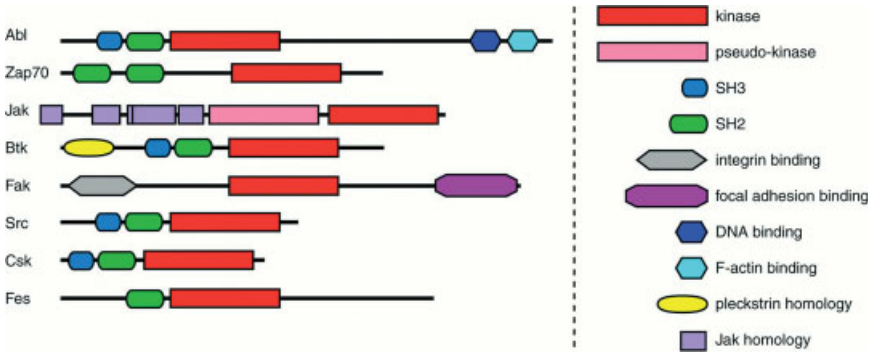


Fig. 8.16 Domain organization for the major subfamilies of nonreceptor tyrosine kinases. The amino terminus is on the left and the carboxy terminus on the right. The lenghts are only approximately to scale. Details of the nonreceptor tyrosine kinases: Csk, Src, Abl: Section 8.3.2; Zap 70: Section 11.3.2; Jak: Section 11.2.2; Fak: Section 11.4. Btk: Bruton kinase.

performed in the immediate vicinity of the cell membrane, whether a signal is received from an activated membrane receptor or is passed on to a membrane-associated protein. To facilitate membrane association, many nonreceptor tyrosine kinases contain N-terminal lipid anchors.

In Fig. 8.16, the structures of the major subfamilies of the nonreceptor tyrosine kinases are shown schematically (review: Hubbard and Till, 2000). In addition to the catalytic domain, the nonreceptor tyrosine kinases often have SH2 and SH3 domains responsible for specific association with substrate proteins. Other sequence motifs mediate the association of the nonreceptor tyrosine kinases with specific subcellular sites and with subcellular structures.

8.3.2
Src Tyrosine Kinase and Abl Tyrosine Kinase

Two of the nonreceptor tyrosine kinases are highlighted here: Src kinase and Abl kinase.

Src Kinase

Src kinase belongs to a family of closely related tyrosine-specific protein kinases involved in the regulation of cell division, cell differentiation, and cell aggregation. At least 9 different protein kinases are numbered amongst the family of Src kinases. Src kinase is involved, for example, in ion channel regulation and in signal transduction via growth factor receptors, integrins and immuno receptors. Functional interactions have been described, e. g., with Tyr-P residues of PDGF receptor, EGF receptor, focal adhesion kinase FAK (see Chapter 11), and scaffolding proteins of the *N*-methyl-D-aspartate receptor complex (review: Parsons and Parsons, 1997).

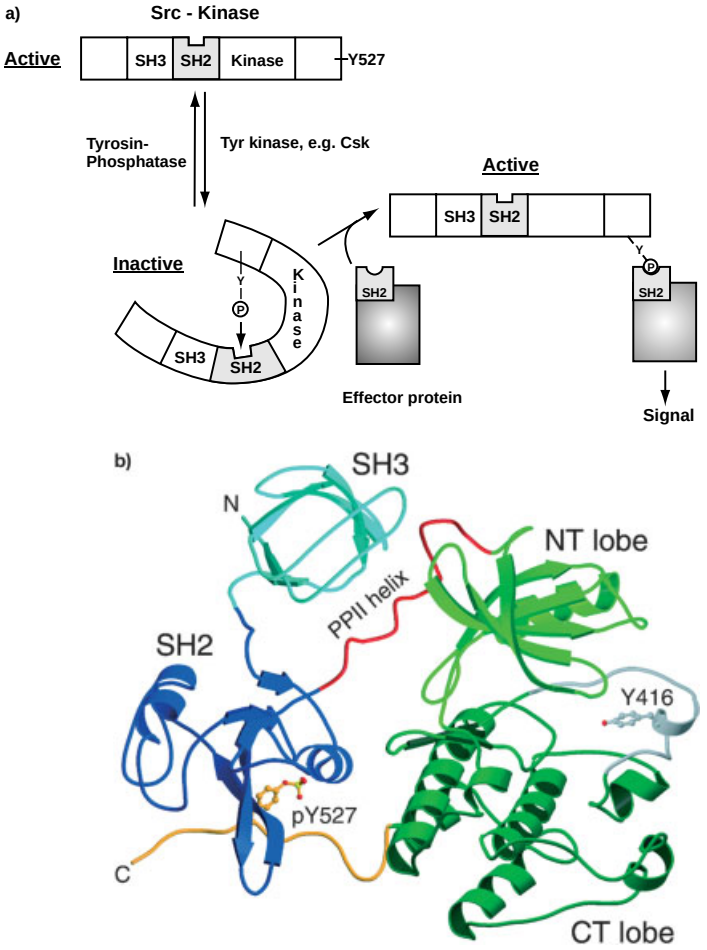


Fig. 8.17 a) Model of regulation of Src kinase by phosphorylation. At the C-terminal end of the Src tyrosine kinase, a Tyr phosphorylation site (Tyr527) is located, which, when phosphorylated by a tyrosine kinase leads to inactivation of the Src kinase. P-Tyr527 binds intramolecularly to the SH2 domain, blocking the kinase activity. Removal of the Tyr phosphate P-Tyr527 by a tyrosine phosphatase converts the Src kinase into the active state again. Activation of the Src kinase can also be brought about by a SH2-containing effector protein; the SH2 domain of this effector protein competes with the SH2 domain of Src for binding to P-Tyr527. Alternatively, P-Tyr527 may bind to the SH2 domain of another signal protein (not shown in the figure). The src kinase is also regulated by phsohorylation . b) Structure of c-Src kinase phosphorylated at Tyrosine 527. Ribbon diagram showing the structure and organization of the "closed conformation" of c-Src kinase. Two aspects of the structure are important for the regulation of c-Src kinase: i) The phosphorylated Tyr 527 of the C-terminal tail is engaged in an intramolecular interaction with the SH2 domain. ii) The SH3 domain binds to the linker between the SH2 domain and the kinase domain. Both interactions are assumed to fix an inactive state of the kinase. A disordered section of the activation segment is shown as a dashed line. The amino- and carboxy-terminal kinase lobes are shown in *light* and *dark green*, respectively. The activation loop in the kinase domain, containing Tyr416, is shown in *gray*. The SH2 and SH3 domains are shown in *dark blue* and *cyan*, respectively. The SH2-kinase linker, which contains a short stretch of polyproline type II helix ((PPII helix) is shown in *red*. The carboxy-terminal tail, which contains pTyr527, is shown in orange.

Src kinase was discovered during the search for the tumor-causing principle of retroviruses. The viral oncogene product of these viruses, v-Src kinase, was the first tyrosine kinase to be identified. In comparison to its cellular counterpart, c-Src kinase, v-Src lacks inhibitory elements that control protein kinase activity. As a consequence of this loss, v-Src kinase is constitutively active and is a potent transforming protein.

These observations and mutation studies have shown that the activity of c-Src kinase is subject to strict regulation. The domain structure of c-Src kinase is shown in Fig. 8.17a. Src kinase carries a myristinic acid residue as a membrane anchor and harbors an SH2 and an SH3 domain N-terminal to its kinase domain. Furthermore, Src kinase possesses two important regulatory phosphorylation sites, namely Tyr 416 in the activation loop and Tyr 527 near the C-terminus.

Numerous studies have demonstrated that the SH2 and SH3 domains as well as phosphorylation of Tyr 527 are implicated in negative control of Src kinase activity. Mutations in the SH2 and SH3 domains that disrupt binding to phosphotyrosine and proline-rich sequences, respectively, stimulate Src activity. Similarly, substitution of Tyr 527 by phenylalanine leads to activation of Src. Based on these observations, a model is now generally accepted according to which Src kinase is maintained in an inactive state by intramolecular interactions involving the SH and SH3 domain and phosphotyrosine 527 (see Fig. 7.17a). According to this model, active forms of Src kinase can be inactivated via phosphorylation at Tyr 527, whereby intramolecular binding of phosphotyrosine 527 to the SH2 domain locks the enzyme in the inactive, autoinhibited state. The repressive phosphorylation at Tyr 527 is catalyzed by the c-Src-specific protein kinase named Csk.

Escape from the inactive state may be achieved in several ways. For example, dephosphorylation of Tyr-P 527 by tyrosine phosphatases will relieve autoinhibition as will binding of the SH2 domain to high-affinity phosphotyrosine residues located on other signaling proteins. In the same way, binding of Tyr-P 527 to high-affinity SH2 domains on other signaling proteins will break up the intramolecular interaction (see below). The importance of phosphorylation at Tyr 527 is underscored by v-Src kinase, an oncogenic variant of Src kinase that is a product of the Rous sarcoma virus. Owing to a C-terminal truncation, v-Src kinase lacks the regulatory site Tyr 527 and is constitutively active, resulting in uncontrolled growth of infected cells.

The molecular mechanisms underlying repression of Src kinase activity have been impressively elucidated by the crystal structure of the inactive form of Src kinase (Fig. 8.17b), in which Tyr527 is phosphorylated (Xu et al., 1997; review:Hubbard and Till, 2000). The crystal structure clearly shows the domain architecture and the interaction between the domains of Src kinase. As already predicted from mutation experiments, Tyr527-P enters into intramolecular binding with the SH2 domain. Unexpectedly, the SH3 domain is involved in an interaction with the linker between the catalytic domain and the SH2 domain, and it helps to fix the SH2 domain in a conformation favorable for interaction with Tyr527-P. These intramolecular interactions stabilize a catalytically unfavorable position of the catalytic residues in the kinase domain and prevent productive ATP binding. A critical structural element for negative regulation is the C-helix of the catalytic domain, which interacts with residues of the linker region and is thereby misaligned when compared to active structures of other

active protein kinases. This inactive conformation of the catalytic domain is very similar to that of CDK2 (see Chapter 13).

Activation of Src kinase requires the disruption of the intramolecular constraints that are mediated by the SH2 and SH3 domains. Through binding of high-affinity ligands, the SH2 and SH3 domains can be displaced from the phosphotyrosine 527 and the linker region, respectively. For example, activation of Src kinase is achieved through SH2 binding to the autophosphorylated PDGF receptor or through SH3 domain binding of the HIV protein Nef.

Another element essential for full activation of Src is Tyr 416 of the activation loop. In the repressed state, this loop is positioned in a way that occludes the substrate-binding site. It is assumed that autophosphorylation of Tyr416 stabilizes an active conformation of the loop.

Abl Tyrosine Kinase

Much of the interest in Abelson (Abl) tyrosine kinase (review: Smith and Mayer, 2002) stems from its involvement in oncogenesis in rodents and in humans. Like many other nonreceptor tyrosine kinases, Abl tyrosine kinase may be converted by mutations into a dominant oncoprotein and may thus contribute to tumor formation. The wild-type form of the Abelson kinase is termed c-Abl; the viral, oncogenic form is termed v-Abl. This mutated enzyme was first discovered as the oncogene of murine Abelson leukemia virus. Apart from the v-Abl enzymes, other oncogenic forms of the Abelson kinase exist. Chronic myelogenic leukemia in humans is caused by a chromosome translocation in which a fusion protein is created from Abl tyrosine kinase and a Bcr protein (cf. Chapter 14). The result is a greatly increased tyrosine kinase activity with very different regulatory properties, to which a causal role in the occurrence of this leukemia is attributed.

The complex structure is a distinctive feature of Abl tyrosine kinase (Fig. 8.18). The enzyme possesses a Tyr kinase domain, an SH2 and an SH3 domain, three nuclear localization signals, a DNA-binding domain, binding domains for microfilament proteins, namely G-actin and F-actin, and a C-terminal myristinic acid residue as a membrane anchor. It is autophosphorylated and phosphorylated by other protein kinases. c-

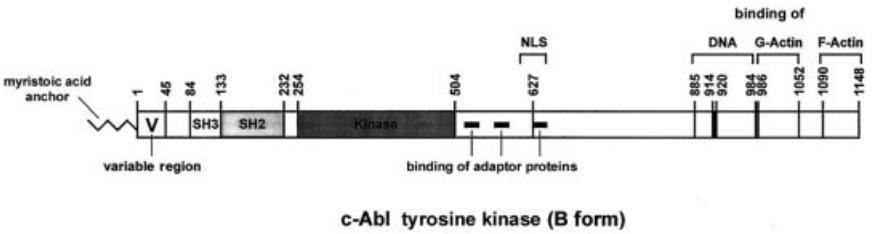


Fig. 8.18 Domain structure of Abl tyrosine kinase. The functionally characterized domains of Abl tyrosine kinase are shown in linear configuration. NLS, nuclear localization signal.

Fig. 8.16. Domain structure of protein tyrosine phosphatases. Linear representation of functional domains of the transmembrane tyrosine phosphatase CD45 and some cytoplasmic tyrosine phosphatases.

Abl tyrosine kinase is an unusual protein kinase in that a large part of it is found in the nucleus, where it interacts via its DNA-binding domain in a rather nonspecific way with DNA and DNA-bound protein complexes. The functions of c-Abl in the nucleus appear to be manifold. It has been implicated, e. g., in programmed cell death, transcription regulation, DNA damage checkpoints and in cell cycle control. The C-terminal domain (CTD, see Section 1.4.3.4) of RNA polymerase II is a well-established substrate of c-Abl in the nucleus. Through a direct interaction with the retinoblastoma protein pRB, c-Abl can be recruited to the E2F-DNA complex (see Section 13.4.2). In this state, the activity of c-Abl is inhibited and pRb phosphorylation at the G1/S transition of the cell cycle is required to relieve inhibition. Activation of nuclear c-Abl is also observed under conditions of DNA damage, and an interaction between c-Abl and the ATM kinase, which is part of a DNA damage-sensing pathway (see Section 13.8), has been shown. Cytoplasmic Abl kinase is found in association with the plasma membrane and with the cytoskeleton, and activation via growth factor receptor signaling pathways has been reported.

8.4 Protein Tyrosine Phosphatases

Protein tyrosine phosphatases play a crucial role in the control of the activity of receptor tyrosine kinases, nonreceptor tyrosine kinases, and the signaling pathways that they regulate. The importance of the tyrosine phosphatases for receptor tyrosine kinase signaling is illustrated by the observation that virtually all receptor tyrosine kinases can be activated, even in the absence of ligand, by treatment of cells with tyrosine phosphatase inhibitors, demonstrating that the activity of tyrosine kinases is continuously controlled by inhibitory tyrosine phosphatase action. As outlined above, the activity of most receptor tyrosine kinases is positively controlled by Tyr-phosphorylation in the activation loop. Protein tyrosine phosphatases that remove these stimulatory phosphate residues will inhibit receptor activity and the biological responses mediated by Tyr-phosphorylation-dependent signaling pathways.

The protein tyrosine phosphatases are therefore elementary components of many signal transduction pathways and, as such, are involved in the control of cell-cell interactions, in signal transduction via growth factors, and in the regulation of the cell cycle (review: Neel and Tonks, 1997; Neel and Tonks, 2001 Östman and Böhmer, 2001). The biological importance of the protein tyrosine phosphatases is underlined by the observation that defects in their activity can lead to phenotypically demonstrable errors in function in higher eucaryotes. One example is the “moth-eaten” mutation in mouse, which is due to a defect in protein tyrosine phosphatase PTP1C. An error in splicing of the gene for the PTP1C protein leads to immunodeficiency and autoimmune disease in the mouse.

Other experiments have demonstrated the biological importance of tyrosine phosphatases in cell signaling as well. Knock-out mice with a disruption of the gene for PTP-1b show increased insulin sensitivity and tissue-specific changes in the phosphorylation status of the insulin receptor.

A medically important protein tyrosine phosphatase is found in the bacterium *Yersinia pestis*, the causative organism of plague. *Yersinia pestis* possesses a highly active protein tyrosine phosphatase which makes an important contribution to the pathenogenicity of this bacterium. The pathogen brings a protein tyrosine phosphatase into the host organism, and this changes the steady state level of tyrosine phosphorylation and leads to extensive disturbance of cellular functions.

8.4.1 Structure and Classification of Protein Tyrosine Phosphatases

The protein tyrosine phosphatases can be roughly divided into two groups: receptor protein tyrosine phosphatases (also called receptor-like protein tyrosine phosphatases) and cytoplasmic protein tyrosine phosphatases. Both groups catalyze the hydrolysis of Tyr phosphate by a common mechanism, and, correspondingly, both groups have a homologous catalytic domain. In addition to these two large groups, there are also tyrosine phosphatases with double specificity that hydrolyze both Tyr-P and also Ser-P. One example is CDC25 phosphatase (see Chapter 13). This group of enzymes is only slightly related to the monospecific Tyr phosphatases.

Receptor Protein Tyrosine Phosphatases

The receptor-like protein tyrosine phosphatases have a transmembrane and, in some cases, a large extracellular domain with a very variable structure (Fig. 8.19). Many, but not all, membrane protein tyrosine phosphatases have two catalytic domains in the cytoplasmic region. The overall structure is very similar to the structure of transmembrane receptors. Only recently, cellular ligands for receptor tyrosine phosphatases have been identified that function as their specific regulators. The receptor tyrosine phosphatase ζ has been found to be specifically inhibited by pleiotrophin, which is a cytokine implicated in tumor angiogenesis. Several studies have demonstrated a role for receptor PTPs in neuronal cell adhesion signaling pathways. In cells of the neural tissue, a surface protein, contactin, has been identified as an extracellular ligand of PTPα.

Cytoplasmic Protein Tyrosine Phosphatases

Cytoplasmically localized protein tyrosine phosphatases have a catalytic domain and other structural elements that specify the subcellular localization and association with effector molecules. These structural elements contain sequence signals for nuclear localization, for membrane association, and for association with the cytoskeleton (see Fig. 8.19) and other signaling proteins. Specifically, the SH2 domains of PTPs have been shown to mediate association with tyrosine phosphates of activated receptor tyrosine kinases. Well-studied SH2-containing PTPs are SH-PTP1, which regulates signaling by hematopoetic receptors, and SH-PTP2 (also known as SYP), which is involved in signaling by growth factor receptors and cytokines (see below, Fig. 8.23).

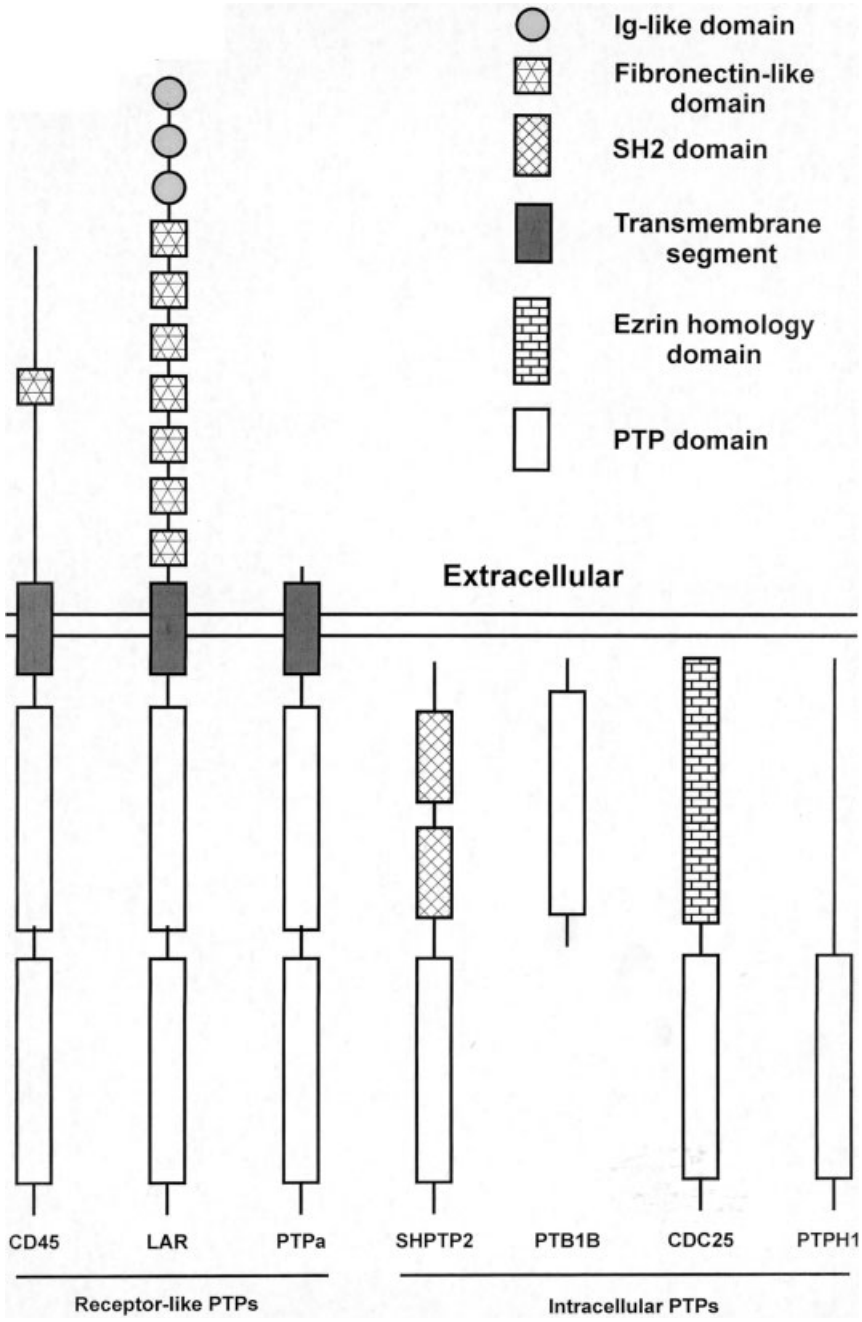


Fig. 8.19 Domain organization of receptor-like and intracellular protein tyrosine phosphatases.

Catalytic mechanism of PTPs

The catalytic center of the protein tyrosine phosphatases includes ca. 230 amino acids and contains the conserved sequence motif H/V-C-(X)5-R-S/T-G/A/P (X is any amino acid), which is involved in phosphate binding and in catalysis. The available structural data on the catalytic domains of protein tyrosine phosphatases support the mechanism shown schematically in Fig. 8.20 (review: Li and Dixon, 2000; Kolmodin and Aqvist, 2001). The invariant Cys and Arg residues of the P loop have a central function in binding and cleavage of the phosphate residue. Another catalytically important residue is an aspartic acid located in a loop distant from the active site cysteine. This loop undergoes a dramatic conformational change when substrate binds to the enzyme.

The Cys residue exists as a thiolate, which carries out a nucleophilic attack on the phosphate of the phosphotyrosine residue. The thiolate is stabilized by the positively charged arginine residue. The Tyr residue is displaced by the thiolate via an "in-line" attack, and an enzyme-bound Cys-phosphate is formed. Discharge of tyrosine from Tyr phosphate is facilitated by the essential Asp residue, which functions as a general acid and donates a proton to the leaving group Tyr phenolate. The phosphate anion in Cys-phosphate is stabilized by the Arg residue and by the positive dipole end of a

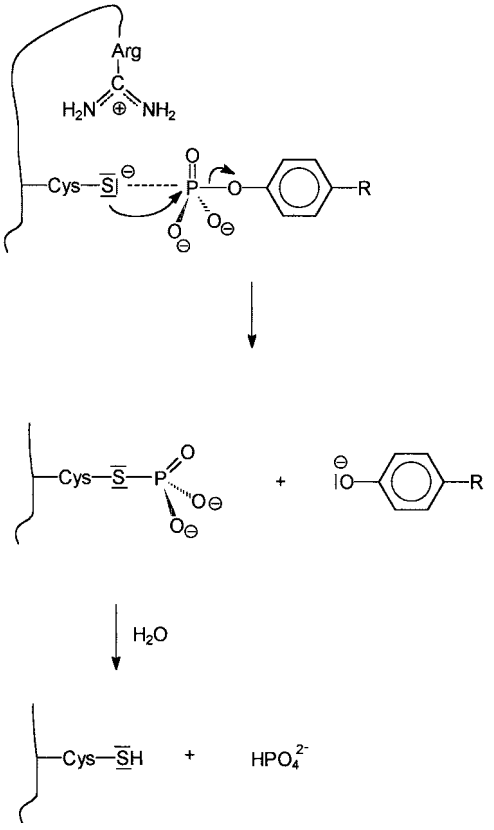


Fig. 8.20 Mechanism of hydrolysis of phosphotyrosine residues by tyrosine phosphatases. Cleavage of phosphate from phosphotyrosine residues takes place by an "in-line" attack of a nucleophilic cysteine thiolate of the tyrosine phosphatase at the phosphate of the phosphotyrosine residue. The negative charge on the thiolate is stabilized by the positive charge of a conserved Arg residue. In the course of the reaction, an enzyme-Cys-phosphate intermediate is formed, which is hydrolytically cleaved to phosphate and enzyme-Cys-SH. The figure shows selected interactions. Other interactions in the active center involved in substrate binding and catalysis are not shown. R: substrate protein.

neighboring α-helix. The phosphate is then released from the intermediate Cys-phosphate by nucleophilic attack of a water molecule.

8.4.2 Cooperation of Protein Tyrosine Phosphatases and Protein Tyrosine Kinases

The cellular functions of protein tyrosine phosphatases are closely associated with signal transduction via protein tyrosine kinases. The growth- and differentiation-promoting signals mediated by receptor tyrosine kinases include receptor autophosphorylations and phosphorylation of effector proteins (see Section 8.1.4). As already outlined for Src kinase, signaling by nonreceptor tyrosine kinases is controlled by inhibitory phosphorylations, among others. According to current ideas, the activity of protein tyrosine phosphatases may have a negative or positive influence on signal transduction via both classes of protein tyrosine kinases. On the one hand, protein tyrosine phosphatases may have an *antagonistic* effect on the activity of protein tyrosine kinases; on the other hand, they may *positively cooperate* with signal transduction via protein tyrosine kinases (review: Östman and Böhmer, 2001).

Negative Regulation of Protein Tyrosine Kinases by PTPs

A schematic representation of how protein tyrosine phosphatases influence signal transduction via protein tyrosine kinases in a negative way is shown in Fig. 8.21a.

An elaborate interplay exists in the cell between the activity of PTPs and receptor tyrosine kinases. Generally, the activity of receptor tyrosine kinases is downregulated by the activity of PTPs. When cells are treated with inhibitors of PTPs, a ligand-independent activation of RTKs and activation of downstream signaling paths occurs, demonstrating the importance of PTPS for maintaining ligand-independent RTK signaling at low levels. In the presence of their ligand, signaling by RTKs is also antagonized by PTPs. Ligand-activated RTKs are rapidly dephosphorylated after inhibition of their tyrosine kinase activity by specific inhibitors. A damping effect of protein tyrosine phosphatases on signaling by activated RTKs may occur, for example, via cleavage of a phosphate residue from an activated receptor tyrosine kinase that has undergone autophosphorylation following ligand binding. Other potential targets of PTPs in controlling RTK signaling are phosphorylated effector molecules of RTKs like, e. g., the insulin receptor substrate (IRS).

There are numerous examples showing that misregulation of tyrosine kinases plays a decisive role in tumor formation and that mutations of the genes of tyrosine kinases can convert these into oncogenes. It has therefore always been hypothesized that protein tyrosine phosphatases may play the role of tumor suppressors (see Chapter 15). Loss of the damping function of protein tyrosine phosphatases in signal transduction can bring about an uncoordinated increase in tyrosine phosphorylation and thus uncontrolled growth. An inhibitory activity of protein tyrosine phosphatases on the formation of tumors has been described many times. For example, protein tyrosine phosphatase PTP1 inhibits tumor transformation by the *Neu* oncogene.

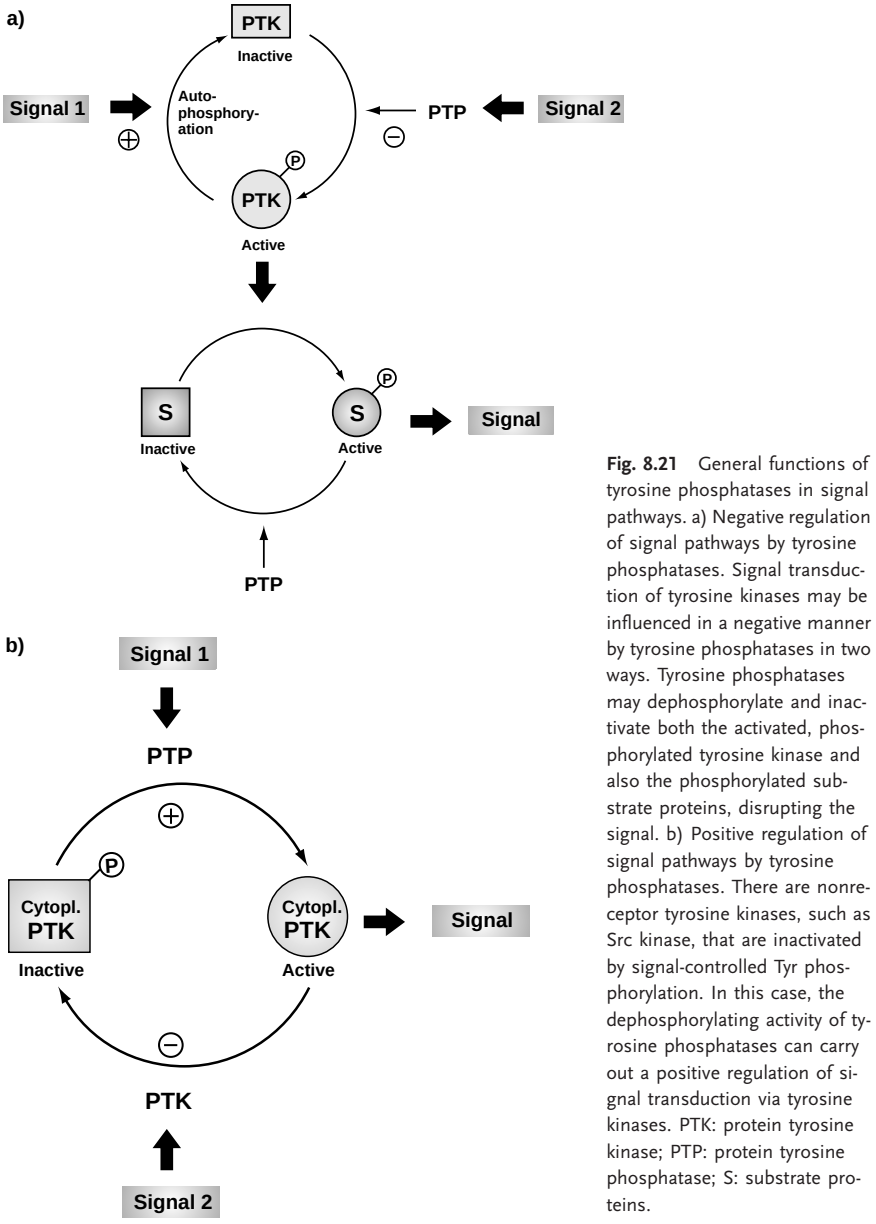


Fig. 8.21 General functions of tyrosine phosphatases in signal pathways. a) Negative regulation of signal pathways by tyrosine phosphatases. Signal transduction of tyrosine kinases may be influenced in a negative manner by tyrosine phosphatases in two ways. Tyrosine phosphatases may dephosphorylate and inactivate both the activated, phosphorylated tyrosine kinase and also the phosphorylated substrate proteins, disrupting the signal. b) Positive regulation of signal pathways by tyrosine phosphatases. There are nonreceptor tyrosine kinases, such as Src kinase, that are inactivated by signal-controlled Tyr phosphorylation. In this case, the dephosphorylating activity of tyrosine phosphatases can carry out a positive regulation of signal transduction via tyrosine kinases. PTK: protein tyrosine kinase; PTP: protein tyrosine phosphatase; S: substrate proteins.

Positive regulation of protein tyrosine kinases by PTPs

Protein tyrosine phosphatases may also carry out a positive regulatory function in a signal transduction by activating protein tyrosine kinases. An example of this regulating mechanism is Src tyrosine kinase. As already explained above (Section 8.2.1), phosphorylation of Tyr527 of Src kinase is linked with inhibition of the kinase activity. The SH2 domain of Src kinase binds in an intramolecular reaction to the Tyr phosphate at the C terminus, leading to blocking of the active center. Activation of Src kinase may be brought about by cleaving off the inhibitory phosphate residue.

Signal transduction via the receptor-like CD45 protein tyrosine phosphatase in cells of the blood forming system requires its intracellularly localized phosphatase activity. The cytoplasmic tyrosine kinases $p56^{lck}$ and $p59^{fyn}$ have been identified as cellular substrates of CD45 phosphatase. Both enzymes are activated by CD45 phosphatase by removal of an inhibitory phosphotyrosine residue. The synergistic action of Tyr phosphatases and Tyr kinases is shown schematically in Fig. 8.21.

8.4.3 Regulation of Protein Tyrosine Phosphatases

Protein tyrosine phosphatases are part of the signaling by receptor tyrosine kinases and, as such, are subject to multiple regulatory influences (Fig. 8.22). The mechanisms are those already highlighted in previous Chapters as central elements of the regulation of activity of signal molecules.

Extracellular Ligand Binding

The activity of receptor tyrosine phosphatases can be positively or negatively regulated by extracellular ligand binding. Interesting information on a possible mechanism of regulation by extracellular ligands was obtained from the crystal structure of the catalytic domain of PTPα, a receptor PTP. PTPα crystallizes as an inactive dimer in which structural elements of one subunit are located in the catalytic site of the other subunit and block the latter (Majeti et al., 1998). The structure suggests that binding of an extracellular ligand promotes oligomerization and inactivation of the receptor-like PTP. This still speculative mechanism involving ligand-induced inhibition is in contrast to the activating effect of ligands on receptor tyrosine kinases.

Oxidation of PTPs

The catalytic cysteine of the PTPs may be reversibly oxidized by different oxidants, including H_2O_2, leading to inactivation of the enzyme. Production of H_2O_2 is induced by a largely unknown mechanism upon stimulation of various receptor tyrosine kinases, including PDGF-R, EGF-R, insulin receptor and FGF-R. In this way, activation of receptor tyrosine kinases may be linked to the inactivation of phosphatase action, allowing for a sharpening of the receptor tyrosine kinase signal.

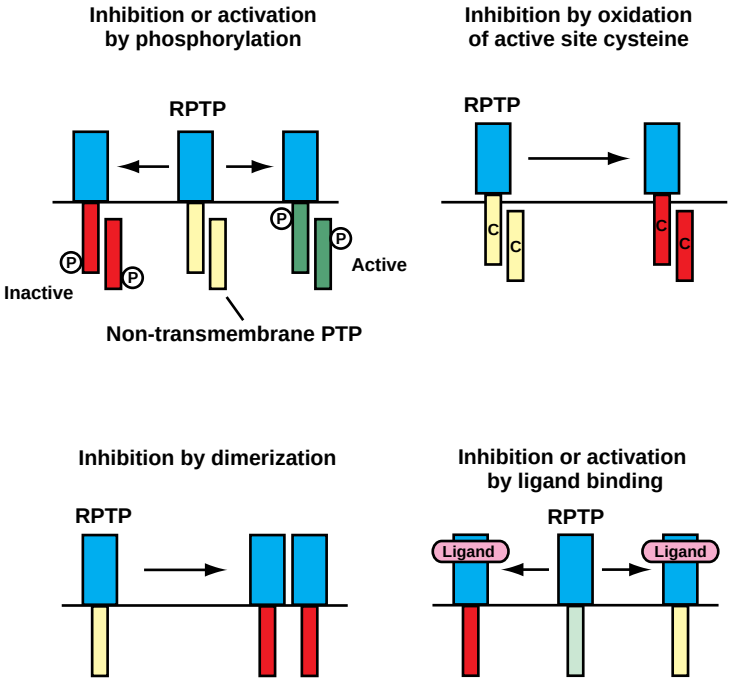


Fig. 8.22 Regulation of protein tyrosine phosphatases (PTPs). The specific activity of PTPs can be either increased or decreased by phosphorylation of serine, threonine or tyrosine residues. Reversible or irreversible oxidation of the active-site cysteine residue (denoted C) also inactivates PTPs. For receptor-like PTPs (RPTPs), dimerization of the catalytic domains has also been proposed as an inhibitory regulatory mechanism. Furthermore, ligand binding to the extracellular domains has been proposed as an inhibitory mechanism. Modulation of specific activity following ligand binding could involve changes in the phosphorylation, oxidation state of the active-site cysteine or dimerization.

Recruitment to phosphorylated RTKs

For the SH2-containing enzymes SH-PTP-1 and SH-PTP-2 (see Fig. 8.23), it has been shown that binding to phosphotyrosine residues on an activated receptor tyrosine kinase is linked to activation of the phosphatase activity. The SH2-phosphotyrosine interaction recruits the phosphatase to the activated receptor and will finally disrupt receptor signaling because of dephosphorylation.

Subcellular Localization

The subcellular localization of protein tyrosine phosphatases is an important aspect of their function. The sequences of cytoplasmic protein tyrosine phosphatases frequently demonstrate sequence signals specifying a particular subcellular localization. This ensures that protein tyrosine phosphatases are only active at defined subcellular sites. Specifically, localization to focal adhesion complexes mediated by proline-rich sequences has been observed for some PTPs. The presence of SH2 domains in cytoplasmic protein tyrosine phosphatases also shows that these are coupled, via

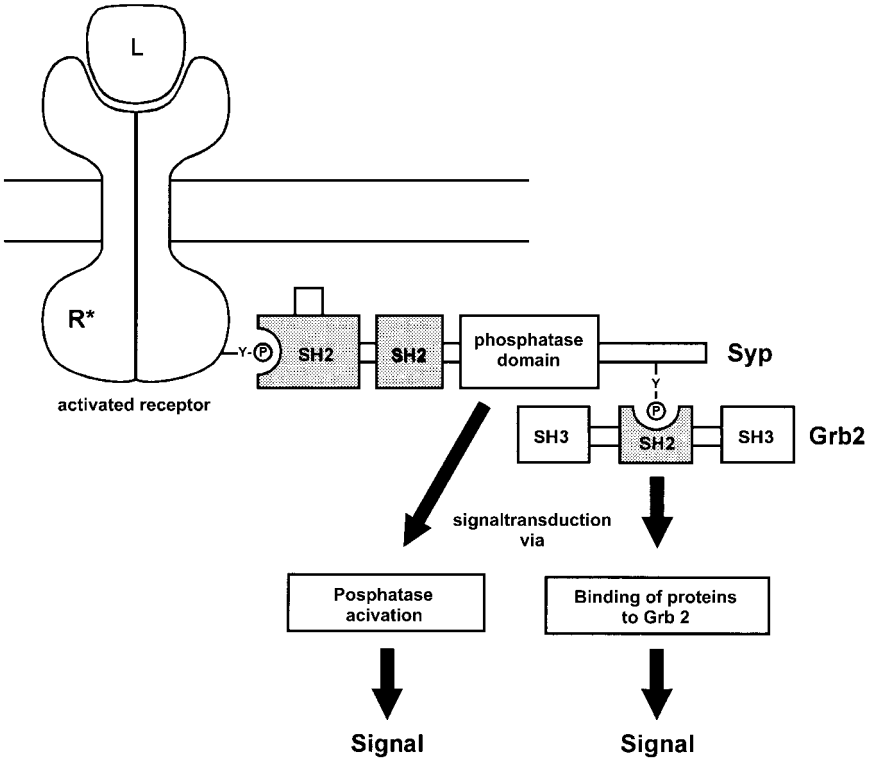


Fig. 8.23 Model of signal transduction via Syp phosphatase. Syp phosphatase has two SH2 domains that mediate binding to the activated, autophosphorylated PDGF receptor. Association of Syp phosphatase with the activated receptor has a twofold effect. The tyrosine kinase activity of the receptor phosphorylates Syp phosphatase at Tyr542. The SH2 domain of the adaptor protein Grb2 binds to the P-Tyr542; Grb2 can bring other signal proteins into the signal transduction process. Furthermore, Syp phosphatase is activated by association with the receptor and can dephosphorylate phosphorylated signal proteins and pass the signal on in this way.

SH2-phosphotyrosine interactions, to specific substrates, where they then perform their actual function.

Ser/Thr Phosphorylation

Another mechanism of regulation of protein tyrosine phosphatases is via Ser/Thr or Tyr phosphorylation. Specific phosphorylation of protein tyrosine phosphatases by Ser/Thr-specific protein kinases of types A and C has been reported (see Neel and Tonks, 1997). This observation indicates the possibility that signal transductions via Ser/Thr kinases and via Tyr kinases/phosphatases may cooperate and that different signal paths may be crosslinked in this way.

Protein tyrosine phosphatases may themselves be the substrate of protein tyrosine kinases and may be phosphorylated at specific Tyr residues. For example, SH-PTP2 (Syp phosphatase) is phosphorylated at Tyr542 on binding to the activated, autopho-

sphorylated PDGF receptor. This phosphorylation site is a consensus binding sequence for the SH2 group of the adaptor protein Grb2 that may may pass the signal to the Ras-MAP kinase pathway (see Chapter 9). SH-PTB2 thus has two distinctly different possibilities to transmit the signal further (Fig. 8.23). On the one hand, the enzyme can involve adaptor molecules in signal transduction; on the other hand, specific substrates can be dephosphorylated.

8.5 Adaptor Molecules of Intracellular Signal Transduction

Signal transduction processes starting from activated transmembrane receptors often have the aim of bringing together particular protein components of the signaling pathway to a defined site in the cell, e. g., to bring an enzyme into the immediate vicinity of its substrate. The cell uses so-called adaptor proteins to bring signal molecules together in a targeted fashion; these adaptor molecules help to decide where and when a certain enzyme, such as a protein kinase, will become active. The adaptor proteins do not have any enzymatic function themselves, but rather they function as a connecting link between different signal proteins, mediating a specific spatial neighborhood in signal conduction (for review see Pawson and Scott, 1997). Furthermore, adaptor proteins serve in targeted subcellular localization of signal events. They are an organizational element in signal conduction, in that they help to assemble multiprotein complexes of signal conduction at specific subcellular sites, enabling spatially concentrated, and thus site-specific, signals to be created. The specificity and regulation of signal conduction are increased, since only certain signal proteins can associate with the adaptor protein. Adaptor proteins provide a scaffold for the assembly of signaling proteins at distinct subcellular sites, and therefore these proteins are also termed scaffolding proteins.

The occurrence of multiple protein modules is characteristic of adaptor proteins; these mediate different protein-protein interactions and can thus bring about cross-linking of signal proteins. Protein modules frequently found in adaptor molecules are SH2 and SH3 domains, PTB domains and PH domains. Furthermore, adaptor proteins may be covalently modified by N-myristoylation and Tyr-phosphorylation. Figure 8.24 shows the schematic composition of some important adaptor molecules.

The occurrence of PH domains and myristoyl modifications suggests that adaptor proteins are involved in particular in the coordination and assembly of signal complexes on the inner side of the cell membrane.

Adaptor proteins are, above all, important elements for controlling the subcellular organization of Tyr and Ser phosphorylation events. Thus, the PTB or SH2 domains of adaptor proteins direct specific interactions with autophosphorylation sites on an activated receptor or with phosphotyrosine residues on other signaling proteins. Furthermore, adaptor proteins like the insulin receptor substrate (see below) often become Tyr-phosphorylated during signaling and thus can provide docking sites for the binding of downstream signaling proteins containing SH2 or PTB domains.

The following adaptor proteins have been shown to be of central importance in cell signaling:

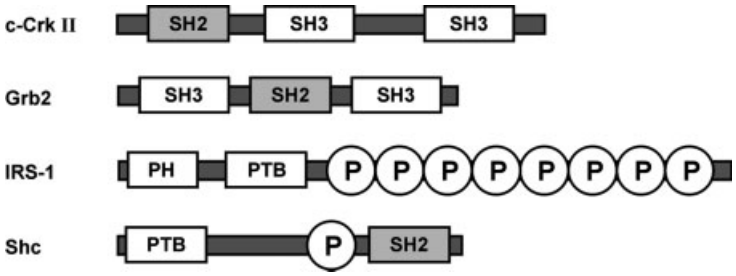


Fig. 8.24 Modular composition of adaptor proteins. Adaptor proteins do not show any enzyme activity of their own, but rather they contain protein modules which help to bind signal proteins into signal pathways. IRS-1: insulin receptor substrate 1; PTB: phosphotyrosine binding domain; PH: pleckstrin homology domain; P: phosphotyrosine-containing binding site for SH2 or PTB domains; HLH: helix-loop-helix DNA binding motif.

- **Grb2**
  The *Grb2 protein* (*Grb: growth factor receptor binding protein*) contains one SH2 domain flanked by two SH3 domains (review: Buday, 1999). It was first identified as a component of signal transduction of growth factors and the Ras signaling pathway (see Chapter 9). The adaptor protein Shc, the EGF receptor, the PDGF receptor and the phosphatase SH-PTP2 have been described as binding partners of the SH2 domain of Grb2 protein. Grb2 protein is tightly bound via its SH3 domain to the Pro-rich domain of the GTP-GDP exchange factor Sos, which can pass the signal by nucleotide exchange to the Ras protein (see Chapter 9). In the form of the Grb2-Sos complex, Grb2 protein functions to generate a coupling between the activated receptor tyrosine kinase and the Ras protein. The membrane association of the Sos protein is necessary for its function as a nucleotide exchange factor (GEF) in the Ras signaling pathway (see Chapter 9). In addition to the Sos protein, other Pro-rich signaling proteins have been known to bind to the SH3 domain of Grb2, indicating a broad spectrum of downstream effector proteins of Grb2 (review: Buday, 1999).
- **Crk**
  The *Crk protein* was first discovered as the transforming principle of the retroviruses CT10 and ASV-1. This family of adaptor proteins seems to be involved in several signaling paths including integrin signaling, growth factor signaling and apoptosis. Mechanistic details of its function are however largely unknown. Crk I has been shown to associate via its SH3 domain with the protein kinase c-Abl and with a nucleotide exchange factor for Ras family members, the C3G protein.
- **IRS**
  The *insulin receptor substrate IRS* is a central adaptor protein that couples the insulin receptor to sequential effector molecules (review: Ogawa et al., 1998). On binding of insulin to the insulin receptor, the tyrosine kinase activity of the receptor is stimulated and the IRS protein binds via its PTB domain to autophosphorylated tyrosine

residues of the receptor. Subsequently the IRS protein is phosphorylated by the activated insulin receptor at several Tyr residues, which then serve as attachment points for sequential effector molecules, e. g., the Grb2-mSos complex, the PI3-kinase and the protein tyrosine phosphatase SHP-2.This function puts the IRS molecules on center stage in insulin action. By recruiting and activating PI3-kinase, a signal in the direction of the Akt kinase pathway is generated, and many of the biological influences of insulin have been linked to activation of this pathway.

- **PDZ-containing adaptor proteins**
  The protein *PSD-95* is an example of a PDZ-containing protein (review: Craven and Bredt, 2000). PSD-95 is found in postsynaptic cells where, via its PDZ domains, it mediates interactions with intracellular domains of receptors such as the *N*-methyl-D-aspartate receptor. The *InaD protein*, which is composed solely of PDZ domains, has an adaptor function in the vision process in *Drosophila* (see Section 8.2.5).

Reference

1 Buday, L. (1999) Membrane-targeting of signalling molecules by SH2/SH3 domain-containing adaptor proteins. *Biochim.Biophys.Acta*, **1422**, 187–204.

2 Cantley, L.C. and Songyang, Z. (1997) Specificity in protein-tyrosine kinase signaling. *Adv.Second Messenger Phosphoprotein Res.*, **31**, 41–48.

3 Craven, S.E. and Bredt, D.S. (2000) Synaptic targeting of the postsynaptic density protein PSD-95 mediated by a tyrosine-based trafficking signal. *J.Biol.Chem.*, **275**, 20045–20051.

4 Durocher, D. and Jackson, S.P. (2002) The FHA domain. *FEBS Lett.*, **513**, 58–66.

5 Ferl, R.J., Manak, M.S. and Reyes, M.F. (2002) The 14–3-3s. *Genome Biol.*, **3**, REVIEWS3010.

6 Gillooly, D.J., Simonsen, A. and Stenmark, H. (2001) Cellular functions of phosphatidylinositol 3-phosphate and FYVE domain proteins. *Biochem.J.*, **355**, 249–258.

7 Heldin, C.H. (1995) Dimerization of cell surface receptors in signal transduction. *Cell*, **80**, 213–223.

8 Hubbard, S.R. and Till, J.H. (2000) Protein tyrosine kinase structure and function. *Annu.Rev.Biochem.*, **69** , 373–398.

9 Hung, A.Y. and Sheng, M. (2002) PDZ domains: structural modules for protein complex assembly. *J.Biol.Chem.*, **277**, 5699–5702.

10 Kolmodin, K. and Aqvist, J. (2001) The catalytic mechanism of protein tyrosine phosphatases revisited. *FEBS Lett.*, **498**, 208–213.

11 Lemmon, M.A. and Ferguson, K.M. (2000) Signal-dependent membrane targeting by pleckstrin homology (PH) domains. *Biochem.J.*, **350 Pt 1**, 1–18.

12 Li, L. and Dixon, J.E. (2000) Form, function, and regulation of protein tyrosine phosphatases and their involvement in human diseases. *Semin.Immunol.*, **12**, 75–84.

13 Macias, M.J., Wiesner, S. and Sudol, M. (2002) WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett.*, **513**, 30–37.

14 Majeti, R., Bilwes, A.M., Noel, J.P., Hunter, T. and Weiss, A. (1998) Dimerization-induced inhibition of receptor protein tyrosine phosphatase function through an inhibitory wedge. *Science*, **279**, 88–91.

15 Margolis, B., Borg, J.P., Straight, S. and Meyer, D. (1999) The function of PTB domain proteins. *Kidney Int.*, **56**, 1230–1237.

16 Neel, B.G. and Tonks, N.K. (1997) Protein tyrosine phosphatases in signal transduction. *Curr.Opin.Cell Biol.*, **9**, 193–204.

17 Ogawa, W., Matozaki, T. and Kasuga, M. (1998) Role of binding proteins to IRS-1 in insulin signalling. *Mol.Cell Biochem.*, **182**, 13–22.

18 Parsons, J.T. and Parsons, S.J. (1997) Src family protein tyrosine kinases: cooperating with growth factor and adhesion signaling pathways. *Curr.Opin.Cell Biol.*, **9**, 187–192.

19 Pawson, T. and Scott, J.D. (1997) Signaling through scaffold, anchoring, and adaptor proteins. *Science*, **278**, 2075–2080.

20 Pawson, T., Gish, G.D. and Nash, P. (2001) SH2 domains, interaction modules and cellular wiring. *Trends Cell Biol.*, **11**, 504–511.

21 Pawson, T., Raina, M. and Nash, P. (2002) Interaction domains: from simple binding events to complex cellular behavior. *FEBS Lett.* , **513**, 2–10.

22 Pawson, T. (2002) Regulation and targets of receptor tyrosine kinases. *Eur.J.Cancer*, **38 Suppl 5**, S3–10.

23 Sanjay, A., Horne, W.C. and Baron, R. (2001) The Cbl family: ubiquitin ligases regulating signaling by tyrosine kinases. *Sci.STKE.* , **2001**, E40.

24 Schlessinger, J. (2000) Cell signaling by receptor tyrosine kinases. *Cell*, **103**, 211–225.

25 Schlessinger J. (2002) Ligand-induced, receptor-mediated dimerization and activation of EGF receptor. *Cell* **110**, 669–72

26 Smith, J.M. and Mayer, B.J. (2002) Abl: mechanisms of regulation and activation. *Front Biosci.*, **7**, d31-d42.

27 Sudol, M. (1998) From Src Homology domains to other signaling modules: proposal of the 'protein recognition code'. *Oncogene*, **17**, 1469–1474.

28 Tonks, N.K. and Neel, B.G. (2001) Combinatorial control of the specificity of protein tyrosine phosphatases. *Curr.Opin.Cell Biol.* , **13**, 182–195.

29 Tsunoda, S., Sierralta, J., Sun, Y., Bodner, R., Suzuki, E., Becker, A., Socolich, M. and Zuker, C.S. (1997) A multivalent PDZ-domain protein assembles signalling complexes in a G-protein-coupled cascade. *Nature*, **388**, 243–249.

30 Yaffe, M.B. and Elia, A.E. (2001) Phosphoserine/threonine-binding domains. *Curr.Opin.Cell Biol.*, **13**, 131–138.

31 Yan, K.S., Kuti, M. and Zhou, M.M. (2002) PTB or not PTB – that is the question. *FEBS Lett.*, **513**, 67–70.

9
Signal Transmission via Ras Proteins

9.1
The Ras Superfamily of Monomeric GTPases

Intracellular signal transduction employs central switching stations that receive, modulate and transmit signals further. Small G-proteins, the Ras protein being a well-known example, make up membrane-associated switching stations of particular importance for growth, differentiation, morphogenesis, vesicular trafficking, and the formation of cytoskeleton. The most prominent and best-characterized member of this class of G-proteins is the Ras protein, which has attracted great interest because of its identification as a human oncogene. Accordingly, the small G-proteins have been classified also as the Ras superfamily of monomeric GTPases. The members of this superfamily are grouped into at least five major subfamilies: the Ras/Rap, the Rho/Rac, the Rab, the Ran and the Arf subfamilies (see Table 9.1). These proteins are monomeric regulatory GTPases of molecular mass 20–40 kDa, which perform a switch function in many biological processes, including intracellular signal transduction.

The principal functions of regulatory GTPases have already been outlined in a general sense in Section 5.4. The members of the Ras superfamily are distinguished from the other regulatory GTPases by their relatively low molecular weight and their occurrence as monomers, but they share the switch properties of the G-proteins: by cycling between the inactive GDP-bound state and the active GTP-bound state, these proteins can receive and transmit signals. In the GTP-bound state, the members of the Ras superfamily interact with downstream effectors. These communicate, in turn, with other signal proteins localized downstream in the signal chain, thereby transmitting the signal further.

Most of the biological functions of the members of the Ras superfamily are linked to the cytoplasmic side of the cell membrane, where specific signals are received and transmitted further. Accordingly, Ras superfamily members contain structural features that mediate membrane association and serve as lipid anchors (see Section 3.7). Prenylation, palmitoylation and myristoylation are post-translational modifications frequently found on Ras superfamily members.

The Ras superfamily members run through the unidirectional GTPase cycle as regulatory GTPases. Their signal-transmitting function can be regulated by three classes of proteins, which have already been discussed in Section 5.4.1 together with the regulatory GTPases:

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

Tab. 9.1 Regulatory GTPases and effector proteins of the Ras superfamily of mammals.

Ras family		GEF, gene or protein name	GAP, gene or protein name
Ras	H-Ras	Ras-GEF, mSos mCDC 25 RasGRF	Ras-GAP; neurofibromin, p120-GAP
	N-Ras		
	Ki-Ras A		
	Ki-Ras B		
	R-Ras		
	M-Ras		
	Rap1A, 1B, 2A, 2B		Rap1-GAP
	RalA, RalB	Ral-GEF	
	TC21 (= k-Rev1)		
	Rit		
Rho/Rac	Rho A, B, C	Dbl	p50 Rho GAP
	Rho 1, 3, 4, 6, 8	Vav1 RhoGEF	p190 Rho GAP
	Rac1, Rac2		
	TC10		
Rab	> 50 different Rab proteins	MSS4 Rab3 GEP	Rab3 GAP
Ran	Ran, TC4	RCC1	Ran GAP1
ARF (ADP-ribosylation factor)	ARF1-6	Sec7	ARF1 GAP

- **GTPase-activating proteins, GAPs**
  The lifetime of the active GTP-bound state may be reduced by *GTPase-activating proteins*. The primary function of the GTPase activating proteins (GAP) is to negatively regulate signal transmission.
- **Guanine nucleotide exchange factors, GEFs**
  The transition from inactive GDP state to active GTP state may be accelerated by proteins that cause the bound GDP to dissociate. The *guanine nucleotide exchange factors* (GEFs) play an essential role in the function of the Ras superfamily members.
- **Inhibitors of guanine nucleotide dissociation, GDIs**
  For the Rab and Rho/Rac families, proteins have been described that bind the GDP form of GTPase and prevent dissociation of GDP. Proteins with these characteristics are known as *guanine nucleotide dissociation inhibitors* (*GDIs*). The GDIs are localized in the cytosol and bind to Rab and Rho/Rac proteins modified with a prenyl residue. Their function is thought to be to extract the Rab and Rho/Rac proteins from the membrane and to prepare a cytosolic pool of these proteins. In this way, GDIs may prevent early dissociation of the bound GDP and premature nucleotide exchange during intracellular translocation of GTPases.

The GAPs and GEFs show specificity for a particular subfamily within the Ras superfamily, and we now know of a large number of different GAPs and GEFs that act on distinct Ras superfamily members (review: Takai et al., 2001). As an example, there are at least nine different GTPase-activating proteins in mammals, which show specificity for the subfamily of Rho proteins. Numerous structural studies on members of the Ras superfamily have shown that the basic mechanisms of nucleotide exchange and GTPase action are well conserved among the different superfamily members, and that the Ras switching station represents a conserved module for signal transduction. Members of the individual subfamilies are distinguished by specific structural insertions or deletions that specify distinct interactions with the cognate GEFs, GAPs and effector proteins, which are highly variable in nature.

A short summary of the properties and biological functions of the five major subfamilies of the Ras superfamily is given below (review: Takai et al., 2001; see also Table 9.1).

- **Ras family**
  This family comprises the Ras proteins in a narrow sense as well as other Ras-related or Ras-like proteins, e. g., the Rap proteins and the Ral proteins (review: Reuther and Der, 2000). Most functions of these proteins are connected to signal transmission across the membrane down to the level of gene expression, thereby influencing proliferation, growth and differentiation of cells. Whereas the regulation and biological function of the genuine Ras proteins have been extensively studied and will be presented in more detail below (see Section 9.2 – 9.8), more limited knowledge exists about the functions of the Ras-related members of the Ras family.
- **Rho family**
  At least 14 members of the Rho family have been identified up to now (reviews:Price and Collard, 2001; Takai et al., 2001). The most prominent members of this family are the Rho proteins, the Rac proteins, and the CDC42 protein. A major function of the Rho/Rac/CDC42 proteins is the regulation of the cytoskeleton in response to extracellular signals in mammalian cells. Such external stimuli are – among other things – growth factors and stress conditions. By influencing, e. g., actin polymerization, the Rho family members relay signals that influence and control essential cellular functions like cell shape change, cell motility, cell adhesion and cytokinesis. Activation of the Rho proteins for signal transmission is usually achieved by the specific GEFs of which a large number have been identified.
  Signals are passed on by the GTP-bound proteins to various downstream effectors, including protein kinases of the MAP kinase cascade (see Chapter 10) and PI3-kinase, to name only a few of these.
- **Rab family**
  Rab proteins exist in all cells and form the largest branch of the Ras superfamily. This family performs a central function in vesicular transport. Rab proteins influence and regulate the budding, targeting, docking and fusion of vesicles as well as processes of exocytosis and endocytosis involving clathrin-coated vesicles. During these functions, Rab proteins cycle between the cytosol and the cell membrane, and this cyle is superimposed on a GDP/GTP cycle. The cytosolic pool of the Rab pro-

teins is thereby maintained in the GDP-bound state by GDI proteins. GEFs and GAPs specific for Rab proteins have been identified as well as various effector proteins (reviews: Segev, 2001; Takai, 2001).

- **Sar/Arf family**
  The Sar and Arf proteins are homologous proteins involved in the transport of specific types of vesicles, namely Cop-coated (Cop, coat protein) vesicles, between the endoplasmatic reticulum and the Golgi apparatus (review: Takai et al., 2001). Similarly to the Rab proteins, the Sar/Arf family members cycle between a cytosolic form and a membrane-associated form regulated by the transition between the GDP- and GTP-bound state, which in turn is controlled by the action of GEFs and GAPs. Membrane translocation is controlled by a myristoyl switch (see Section 3.7.5), where GDP/GTP exchange by GEFs induces a conformational change that allows the myristoylated N-terminal helix of Arf proteins to interact with phospholipid bilayers, thereby promoting membrane insertion (Goldberg, 1998). Examples of effector proteins are the Cop-components of vesicles, among others.
- **Ran family**
  The Ran protein (Ran: Ras-related nuclear protein) is an essential component of nucleocytoplasmic transport. There is only one gene for Ran in human cells. During nucleocytoplasmic transport, the Ran protein interacts in a cyclical manner with various import and export receptors, thereby allowing the transport of cargo proteins in and out of the nucleus. An essential feature of the cyclical transport is the asymmetric distribution of the GDP- and GTP-bound forms of Ran between the nucleus and the cytoplasm, which in turn is caused by an asymmetric distribution of GEFs and GAPs for Ran. For a detailed review of Ran, the reader is referred to Kuersten et al, (2001).

9.2 General Importance of Ras Protein

Within the Ras superfamily of monomeric GTPases, the Ras family historically has attracted the most interest because of its involvement in growth regulation and cell proliferation. The Ras proteins process growth-promoting signals received by receptor tyrosine kinases and by receptors with associated tyrosine kinase activity, and transmit these into the cell interior (Fig. 9.1). Furthermore, Ras family members transduce signals from G protein-coupled receptors and from other monomeric GTPases. Ras protein signaling has now been found to influence not only cell growth, but also cell differentiation, cell morphology, and apoptosis. Most of these effects are mediated through Ras-signaling-induced changes in transcription.

The general importance of Ras proteins in growth regulation was recognized at the beginning of the 1980s, when it was demonstrated that close to 30 % of all solid tumors in humans show a mutation in the Ras gene coding for a 21 kDa product, named the Ras protein or $p21^{ras}$.

Interest in the structure and function of Ras protein and its relatives was kindled in particular because it was established that certain positions in the Ras protein are par-

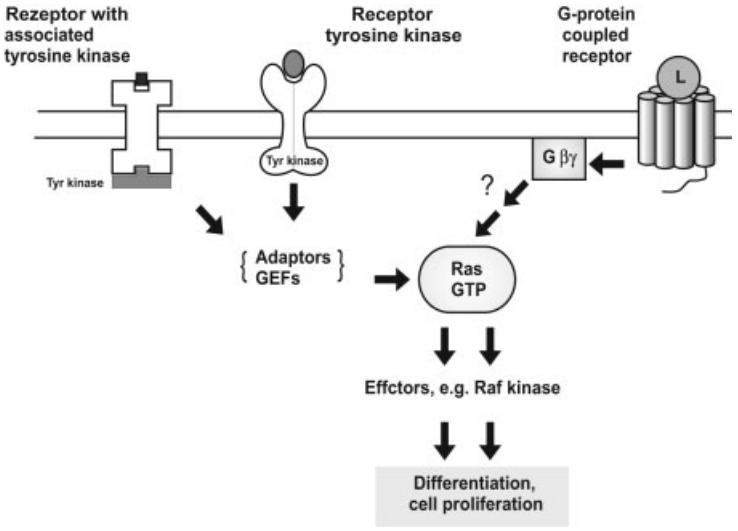


Fig. 9.1 The Ras protein as a central switching station of signaling pathwayss. A main pathway for Ras activation is via receptor tyrosine kinases, which pass the signal on via adaptor proteins and guanine nucleotide exchange factors to the Ras protein. Activation of Ras protein can also be initiated via G-protein-coupled receptors and via transmembrane receptors with associated tyrosine kinase activity. The membrane association of the Ras protein (see Fig. 9.6) is not shown for clarity. In addition, not all signaling pathways that contribute to activation of the Ras protein are shown, nor are all effector reactions. G$\beta\gamma$: $\beta\gamma$-complex of the heterotrimeric G proteins; GAP: GTPase activating protein; GEF: guanine nucleotide exchange factor.

ticularly sensitive to oncogenic mutations. Replacement of Gly12 in the Ras protein with any of the other natural amino acids (except Pro) leads to an increase in the tumor-transforming potential of Ras protein. Its small size and apparently simple function made the Ras protein an ideal subject for investigation of structure-function relationships of a central signal protein. Ras protein is currently the best-characterized signal protein. The extensive structural information available on Ras protein has helped us to understand molecular mechanisms of signal transduction and the cause of tumor-associated misregulation at the molecular level.

The Ras gene and its gene product were first found in retroviruses that trigger sarcoma-type tumors in rats (*Ras* = *rat sarcoma*). Mammals have at least three different Ras genes: ***H**(arvey)*-**ras**, ***K**(irsten)*-**ras*** or ***K-ras***, and the ***N**(euroblastoma)*-**ras*** gene, with the *K-ras* gene producing a major **(K-Ras 4B)** and a minor **(K-Ras 4A)** splice variant. For each of these genes, oncogenic mutations have been found in human tumors. The most frequently mutated gene is *K-ras*, with 70–90% mutations in pancreatic cancer and 20–50% in lung cancer (review: Ellis and Clarke, 2000). It is not yet known to what extent the three gene products, which are highly homologous, have distinct functions. There is experimental evidence that they act at different microdomains of the cell membrane. The proteins coded for by these three genes represent the Ras proteins

in a narrow sense. Most of the structural and biochemical data are available for the H-Ras protein. This is referred to in the following as "*the Ras protein*" for simplicity. It is assumed that the basic knowledge of structure and function of the H-Ras protein also applies to the other Ras proteins.

9.3 Structure and Biochemical Properties of Ras Protein

The Ras protein is a monomeric GTPase of ca. 21 kDa that undergoes the typical G-protein cycle of activation and inactivation, thereby functioning as a switch in signal transmission. The GTP-bound form represents the active, switched-on state; the GDP-bound form is the inactive, switched-off state. The transition between the active and the inactive forms occurs in a unidirectional cycle (see Fig. 9.2). Highly resolved crystal structures exist for both forms of Ras protein and for oncogenic mutants of Ras protein. Furthermore, structural information is available on the binding of a Ras-related protein, the Rap protein, with the effector Raf kinase (Nassar et al., 1995), on the binding of a GAP protein (Scheffzek et al., 1997), and on the binding of a GEF (Boriack-Sjodin et al., 1998). 3D-structures are also available for other members of the Ras superfamily, both alone or in complex with the cognate GEFs or GAPs (reviews: Geyer and Wittinghofer, 1997; Vetter and Wittinghofer, 2001).

The *lifetime* of the active, GTP-bound state of the Ras protein is of great importance for the signal-transducing function of Ras protein. Only in the GTP state can the signal be transmitted to the effector molecule next in sequence. The time window available for signal transduction is determined by the rate of GTP hydrolysis. A low rate of GTP

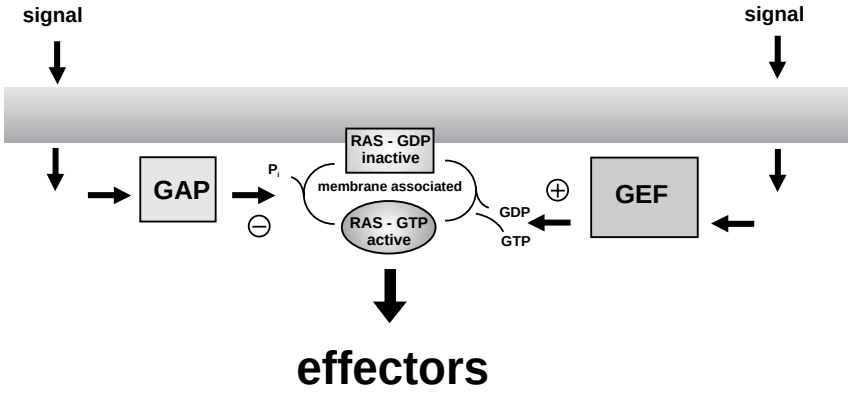


Fig. 9.2 The GTPase cycle of the Ras protein. Conversion of the inactive RasGDP complex into the active RasGTP complex is brought about by guanine nucleotide exchange factors (GEFs). The activated state of the Ras protein is terminated by hydrolysis of the bound GTP. The help of a GTPase actvating protein (GAP) is required, due to the intrinsically slow GTPase activity of the Ras protein. Ras protein performs all its functions in close association with the cell membrane. It carries a membrane anchor and the effector proteins preceding and following in sequence are also associated with the membrane.

hydrolysis and consequently longer occupancy of the active GTP state is associated with a high intensity of signal transduction. Reduction of the time window by stimulation of GTPase leads to weakened signal transduction.

Considered in isolation, the Ras protein is a very inefficient, not to say a "dead" enzyme. On the one hand, the intrinsic rate of GTP hydrolysis is very low, and, on the other hand, the complex of Ras protein and GDP is very stable and only dissociates very slowly. The rate constants of the two processes are in the region of $10^{-4}\,s^{-1}$. Both reactions may be accelerated by Ras-specific GEFs and GAPs in the process of signal transduction, however, and these proteins therefore are essential elements of the switch function of Ras proteins.

Activation of GTPase

Under the influence of Ras-specific GTPase-activating proteins, the rate of GTP hydrolysis of the Ras protein may be increased up to 10^5-fold. The GTPase-activating proteins control the activity state of Ras protein by drastically reducing the lifetime of the active GTP state. Because of this property, they function as negative regulators of the Ras protein.

Acceleration of Nucleotide Exchange

The rate of GDP dissociation is subject to control by *G-nucleotide exchange proteins* (*GEFs*), which promote dissociation of bound GDP and thus function as positive regulators of the Ras protein. The GEFs are the input components of Ras signaling. They receive signals from upstream signaling proteins or protein complexes and transmit these to the Ras protein. By inducing GDP/GTP exchange, the activated GTP-state of Ras proteins is produced, and further transmission of the signal to downstream effectors is now possible.

9.3.1 Structure of the GTP- and GDP-bound Forms of Ras Protein

The structure of the GTP-bound form of the Ras protein is shown in Fig. 9.3.

The Ras protein, as a regulatory GTPase, shows the G-domain typical of the superfamily of regulatory GTPases (see Fig. 5.12). The sequence motives characteristic of regulatory GTPases (cf. Section 5.4.3) are involved in binding the nucleotide and Mg^{2+}. Three structural elements are of particular importance for the switch function of Ras protein: the P-Loop, the switch I region, and the switch II region. All three structural elements contact the γ-phosphate of GTP. Upon GTP hydrolysis, the switch I and switch II regions undergo significant changes in conformation (see below).

The P loop winds around the β and γ phosphates and contributes most of the energy for GTP binding. Gly12 is located in the P loop; this amino acid is most frequently mutated in oncogenic mutants of Ras protein (mutation "hotspot").

The switch I region (residues 30 – 38) and the L2 loop thereof is of particular importance for the biological activity of the Ras protein. An essential function is performed by Thr35 of the L2 loop, which participates in H-bridge contacts to the γ-phos-

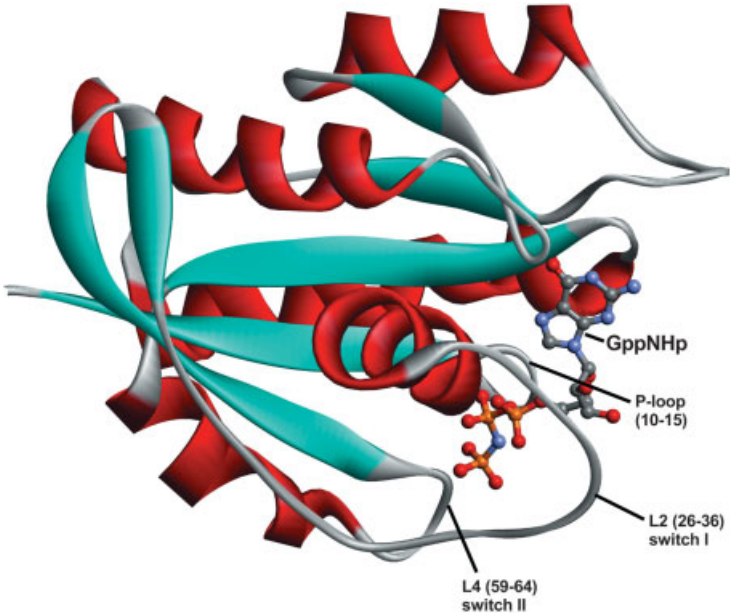


Fig. 9.3 Structure of the GTP form of the Ras protein. Crystal structure of the Ras protein in complex with the GTP analog $\beta\gamma$-imino-GTP(GppNHp). The figure shows the P loop, in which Gly12 is located, and the L2 and L4 loops, which have a switch function in GTP hydrolysis. The numbers give the sequence positions of amino acids in the loops.

phate and to Mg^{2+}, and thus helps to correctly position the GTP-Mg^{2+} complex in the active center. The L2 loop of the switch I region is also known as the *effector loop*. It is an important part of the effector domain of the Ras protein, and signals are received and passed on via this domain.

The switch II region (residues 59–67) includes a conserved DXXG motif. Gly60 of this motif forms an H-bridge to the γ-phosphate. Switch II also contains the catalytically essential Gln61 residue and is involved in the interaction with GTPase-activating proteins.

The structural difference between the active GTP-state and the inactive GDP-state of the Ras protein is primarily confined to the switch I and switch II regions. The conformational change can be described best by a "loaded spring" mechanism (see Fig. 5.22), where the two switch regions are fixed by the γ-phosphate of GTP in a tense state. Upon GTP-hydrolysis, the two switch regions are allowed to relax into the GDP-specific position. Thereby, the coordination of switch I to the γ-phosphate and to Mg^{2+} is lost as well as the interaction of the conserved Gly60 in switch II with the γ-phosphate.

It is not surprising that residues corresponding to switch I and switch II, which define the conformational differences between the inactive GDP form and the active GTP state of Ras, are involved in recognition of the Ras effectors, the immediate downstream components in the Ras signaling pathway (see Sections 9.7 and 9.8). Residues 32–40 comprise the core Ras effector domain, which is essential for all effector interactions.

9.3.2
GTP Hydrolysis: Mechanism and Stimulation by GAP Proteins

In the high-resolution structure of the GTP form of Ras protein, a tightly bound water molecule is visible located in an optimal position for nucleophilic attack on the γ-phosphate. The water molecule is fixed in a defined position by H-bridges with Gln61 and Thr35. As described in Section 5.4.4 for the α-subunits of the heterotrimeric G-proteins, GTP hydrolysis takes place by an "in-line" attack of the nucleophilic water molecule on the γ-phosphate, for which a pentagonal, bipyramidal transition state is postulated.

The rate of GTP hydrolysis in the Ras-GTP complex is very low and would not be suitable for cellular signal transduction, which normally includes complete inactivation within minutes after GTP activation. Therefore, termination of the active GTP state requires the participation of GAPs as an essential step, increasing the rate of GTP hydrolysis by several orders of magnitude. The molecular basis of this stimulation was explained by structural determination of the Ras-GAP complex (Scheffzek et al., 1997). The crystal structure of the complex of a fragment of p120-GAP (GAP-334 in Fig. 9.4) and Ras*GDP*AlF_3 ended a long discussion on the mechanism of GTPase activation (Fig. 9.4).

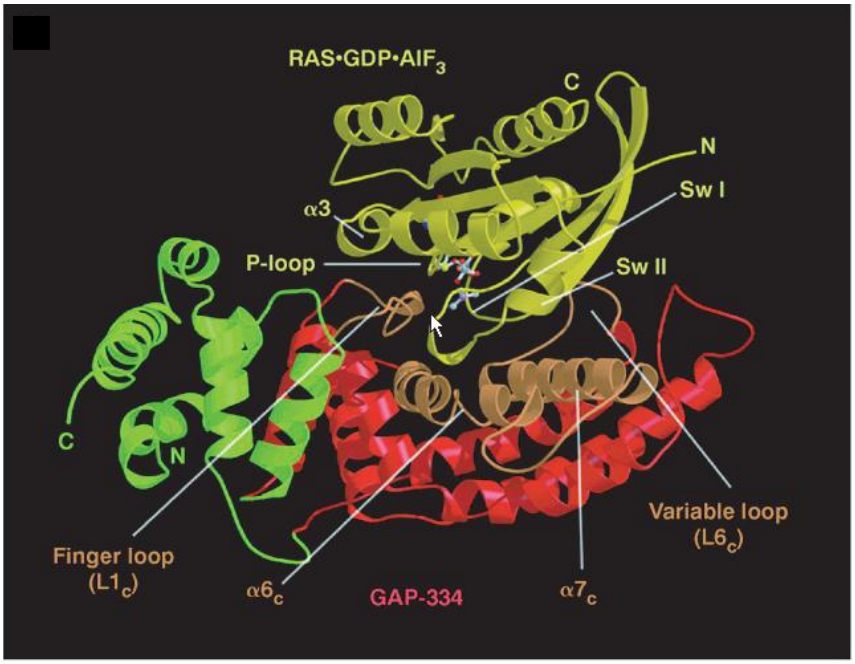


Fig. 9.4 Ribbon diagram of the complex of GAP-334 and the Ras protein. Ras*GDP*AlF3 is shown in yellow. Domains of GAP-334 are shown in green, red and brown.

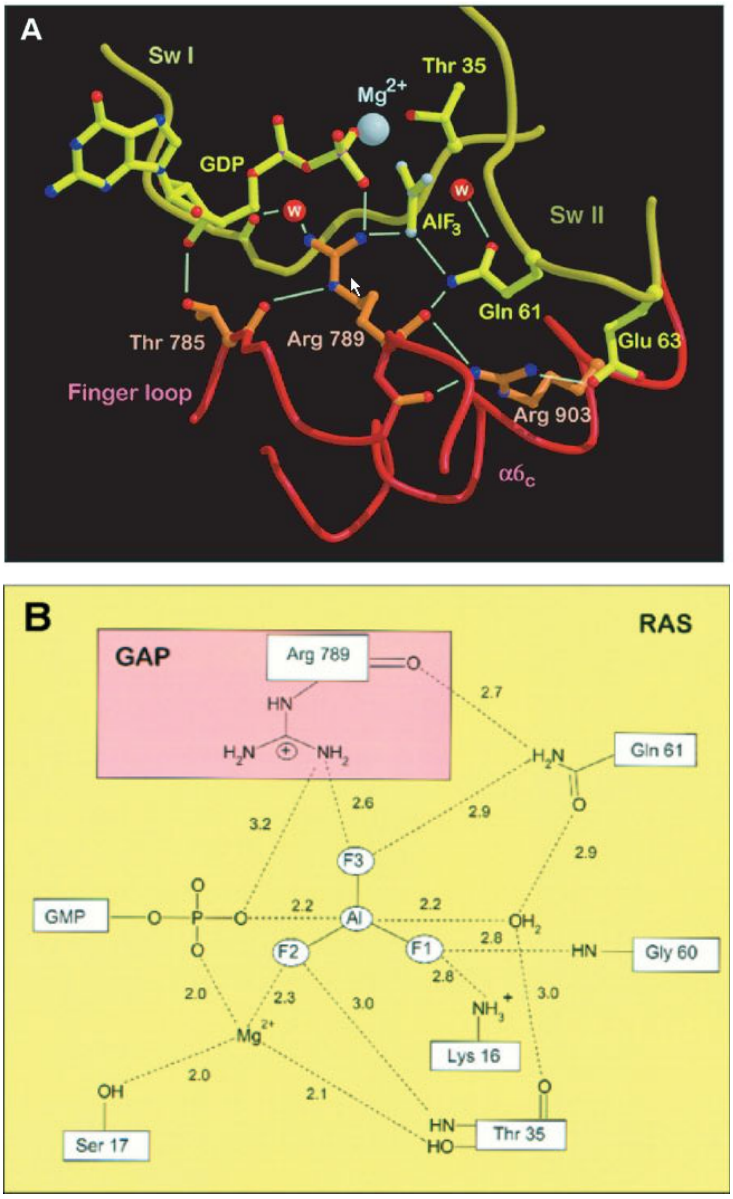


Fig. 9.5 The Ras-GAP complex. A) Structural view of the active site, with the important elements of catalysis.
B) Schematic view of the active site.
C) Cartoon of the mechanism of GTP hydrolysis in the Ras-GAP-GTP complex. Binding of Ras and GAP to GTP induces a shift of negative charges to the oxygens of the β-phosphate. Mg^{2+}, Lys16 of Ras and Arg789 of GAP are the key residues holding like molecular tweezers charges on the β-phosphate. This charge shift reduces the free energy for cleavage of the bond between the β- and γ-phosphate. By this dissociative mechanism of GTP hydrolysis, P_i is strongly bound as an intermediate (B) and is released in a second step (B→C). From Allin et al., 2001.

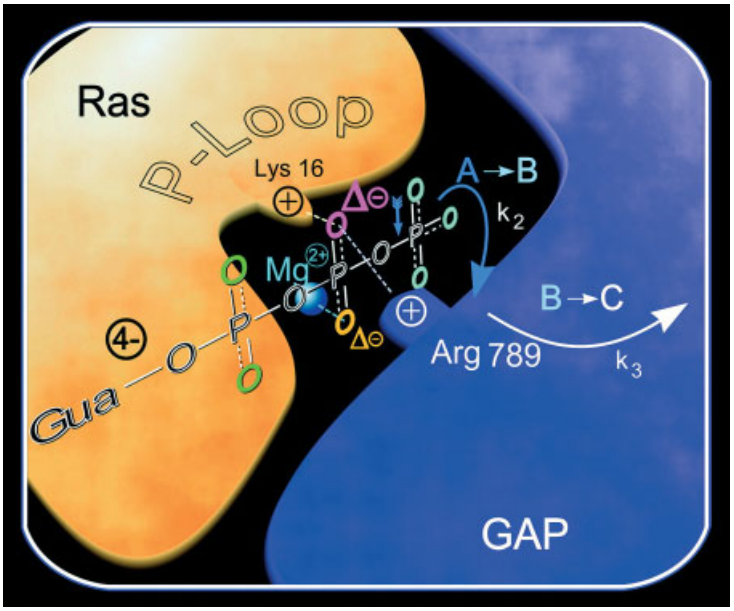


Fig. 9.5 C

The structural data show that the GAP protein actively participates in catalysis by making an Arg residue available, which helps to stabilize the transition state of GTP hydrolysis. In a structural element of the GAP protein known as a finger loop, an invariant Arg residue (R789) is located that interacts with AlF_3; the latter adopts the position of the γ-phosphate in the transition state of GTP hydrolysis (see Section 5.5.4). Next, Arg789 has the role of neutralizing the charge of the γ-phosphate developed in the transition state. Furthermore, Arg789 helps to stabilize the L4 loop of the Ras protein, which is a part of the switch II region. A detailed kinetic analysis of GAP-mediated GTP hydrolysis has shown that GTP hydrolysis proceeds by a "dissociative" mechanism, where GAP binding to Ras-GTP shifts negative charges from the γ- to the β-phosphate and where the γ-phosphate is fixed by Arg789 of GAP (Allin et al., 2001; see Fig. 9.5c). In this process, a central function in GTP hydrolysis is attributed to Gln61 of switch II, since it is located in an ideal position for exact alignment of the water molecule and for stabilization of the transition state of GTP hydrolysis (Fig. 9.5a). The observation that position 61 – after position 12 – is the second most frequent site of oncogenic mutations in solid tumors is in agreement with the central importance of Gln61 for GTP hydrolysis.

The GTPase of the α-subunits of heterotrimeric G-proteins also uses an Arg residue (Arg178 in Fig. 5.20) for stabilization of the transition state of hydrolysis. In contrast to the Ras protein, this is localized in the cis configuration on the α-subunit itself and is found in the linker between the helical domain and the G-domain.

9.3.3
Structure and Biochemical Properties of Transforming Mutants of Ras Protein

The *ras* genes are the most frequently mutated oncogenes detected in human cancer. Comparison of the biochemical properties of mutated Ras proteins with the wild-type Ras protein shows that *oncogenic activity* correlates with increased lifetime of the GTP form. Ras proteins can be converted to oncogenic, transforming forms by mutations in particular at positions 12, 13 and 61. Gly 12 located in the P-loop is especially sensitive to amino acid substitutions, and mutations at this position are the most frequent Ras mutations found in human tumors. Replacement of Gly12 with any amino acid other than proline leads to oncogenic activation of Ras protein. The rate of GTP hydrolysis of oncogenic mutants is about 90 % lower than that of the wild type. It is important that the low rate of GTP hydrolysis cannot be increased by GTPase-activating proteins, in contrast to the wild-type protein. With respect to the lifetime of the activated GTP state, there is therefore a large difference between the wild-type Ras protein and the oncogenic mutated Ras proteins. In the presence of GTPase-activating proteins, the oncogenic mutants of Ras protein spend a ca. 10^5-fold longer period in the activated state than does the wild-type Ras protein and can transmit a dominant signal in the direction of cell proliferation, favoring tumor transformation.

Comparison of the wild-type structure with the structure of oncogenic Ras proteins, in which Gly12 is replaced by other amino acids, shows only small structural changes in the active center. Only when the structure of the Ras-GAP complex was obtained did the means become clear by which oncogenic G12 mutants influence Ras signal transduction. The G12 of the P-loop is located very close to the main chain of the Arg finger of the GAP protein and to the Gln61 of the Ras protein. Replacement of glycine by other amino acids would lead to Van der Waals repulsion and thus to displacement of the Arg finger and of Gln61. In the oncogenic G12 mutant of the Ras protein, an active role of the Arg finger in GTP hydrolysis is, according to this model, no longer possible.

The effect of oncogenic mutations at position 61 of the switch II region can also be explained using the Ras-GAP complex. Gln61 has a central function in GTP hydrolysis in that it contacts and coordinates the hydrolytic water molecule and the O-atom of γ-phosphate of GTP and thus stabilizes the transition state. Amino acids with other side chains apparently cannot fulfill this function, as shown by the oncogenic effect of Gln61 mutants in which Gln61 is replaced by other amino acids (other than Glu).

9.4
Membrane Localization of Ras Protein

The function of the Ras protein in cellular signal transduction is inseparably bound up with the plasma membrane. The Ras proteins associate with the inner side of the cell membrane with the help of lipid anchors, such as farnesyl residues and palmitoyl residues (see Section 3.7).

Farnesylation of the Ras protein occurs at the C-terminal CAAX sequence (A: aliphatic amino acid, X: Ser or Thr). The farnesyl residue is attached, with the help of a farnesyl protein transferase and via a thioether bond, to the Cys residue of the CAAX sequence. Next, the last three amino acids are cleaved off by proteases, and the carboxyl group of the C-terminal cysteine residue undergoes a methylesterification (Fig. 9.6). In addition, the Ras proteins have a palmitinic acid anchor at different Cys residues in the vicinity of the C terminus. The membrane localization of the K-Ras protein is also supported by a polybasic sequence close to the C terminus (see Section 3.7).

Membrane anchoring via the C-terminal modifications is absolutely necessary for the function of the Ras proteins. The lipid anchors, in contrast, have no influence on the catalytic activity of Ras GTPase. Rather, the membrane anchoring of the Ras protein has the role of bringing the latter to the membrane inner side, into the neighborhood of its upstream signaling partners, e. g., the Grb2-mSos complex that associates with activated receptor tyrosine kinases (see Section 9.6). Deletion of the sequence signals necessary for lipid modification in Ras protein leads to loss of the signal molecule function. Because of the essential function of lipid modification for signal trans-

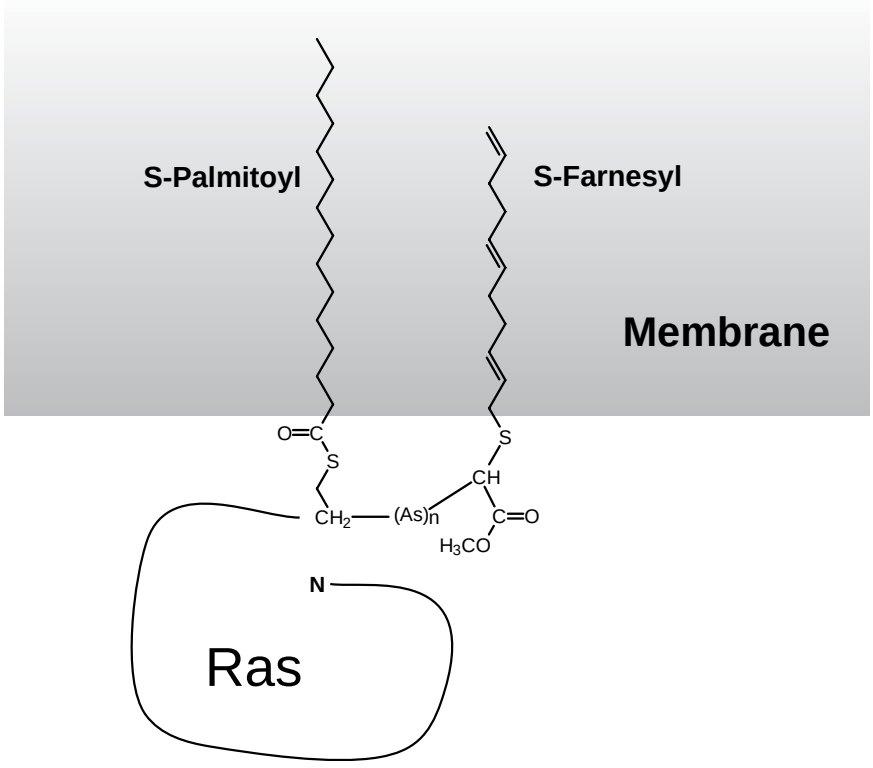


Fig. 9.6 Lipid anchor of the Ras protein. Membrane association of the Ras protein is mediated via a palmitoyl and a farnesyl anchor.

duction via Ras proteins, great efforts are being made to develop inhibitors of farnesylation and to employ these therapeutically as inhibitors of signal transduction via Ras protein in tumors (review: Garbay, 2000).

9.5 GTPase-activating Protein (GAP) in Ras Signal Transduction

As outlined above, GAP proteins are essential components of signal transmission by Ras. The GAP proteins specific for the Ras proteins are known as *Ras-GAP proteins*. The importance of Ras-GAP proteins for signal transduction via Ras proteins is shown by oncogenic Ras mutants with amino acid substitutions at positions 12, 13 and 61, which are resistant to the influence of GAP and show constitutive activation of Ras signal transduction.

Five Ras-GAP proteins are known in mammals, of which the p120-GAP protein and neurofibromin (also abbreviated as NF1) are the best characterized. The domain structure of p120-GAP is shown in Fig. 9.7. p120-GAP has a hydrophobic amino terminus, two SH2 domains, an SH3 domain, a pleckstrin homology domain, and a domain that is homologous to the calcium-binding domain of phospholipase A2. The catalytic domain for GAP activity is found in a 250-amino-acid section close to the C terminus; three other highly conserved sequence elements are also found in this region.

The gene for the GAP protein neurofibromin is deleted in *Recklinghausen neurofibromatosis type I* disease. The protein shows a high degree of homology with the yeast GAP proteins IRA1 and IRA2, and it may complement loss of the IRA functions in yeast. The homology with p120-GAP is much lower and is limited to the catalytic domain. The stimulating influence of neurofibromin on GTPase activity of Ras protein is comparable to that of p120-GAP. The importance of neurofibromin for growth regulation is emphasized by the observation that mutations leading to loss of catalytic activity have been found in the neurofibromin gene in neurofibromatosis patients.

The GAP proteins stimulate GTPase activity of the corresponding Ras protein by an active role in catalysis, and the function of negative regulation of Ras signal transduction is generally attributed to them. GAP proteins control the intensity of signal transduction via Ras proteins by reducing the lifetime of the active state of the Ras protein and thus reducing the number of Ras proteins in the GTP state.

Fig. 9.7 Domain structure of p120-GAP. The functional domains of p120-GAP are shown in linear form. PH; pleckstrin homology domain; SH: Sarc homology domain; A2: possibly Ca^{2+}-dependent phospholipid binding motif; 1,2,3, conserved sequences characteristic for GTPase-stimulating activity.

The presence of SH2 and SH3 domains in p120-GAP indicates a role in signaling pathways starting from receptor tyrosine kinases. In fact, the SH2 domains of p120-GAP mediate specific binding to phosphotyrosine 771 of the β type of PDGF receptor (see also Fig. 8.9). Interactions of GAP proteins with other signaling proteins have also been reported. The precise function of GAP interactions with signaling proteins other than the Ras protein, however, remains to be determined.

9.6 Guanine Nucleotide Exchange Factors (GEFs) in Signal Transduction via Ras Proteins

Ras protein is a central switching station in intracellular signal transduction, which receives, modulates and passes signals on. The Ras protein receives, in particular, signals promoting growth and differentiation, which start from activated receptor tyrosine kinases and other transmembrane receptors and are transmitted through the cell membrane to the Ras protein. Signal transduction between the activated receptors and the Ras protein takes place via *guanine nucleotide exchange factors*, *GEFs*. The GEFs, together with adaptor proteins, form the link between activated receptor tyrosine kinases and Ras protein. The role of the GEFs is to pass the signal to the Ras protein, converting the latter from its inactive GDP form to the active GTP form (see Fig. 9.2).

9.6.1 General Function of GEFs

The question of which signal protein precedes the Ras protein in higher organisms was unclear for a long time. Genetic investigations in *Drosophila* and in the nematode *Caenorhabditis elegans* showed that at least two types of protein are involved in establishing bonding between the activated receptor tyrosine kinase and the switching station of the Ras protein. These are SH2/SH3-containing adaptor proteins (see Section 8.5) and nucleotide exchange factors. Corresponding proteins were then identified in mammals. For mammals, the adaptor protein in question is the *Grb2 protein* (Grb=growth factor receptor binding protein). The Grb2 protein exists in the cell in a tight complex with the nucleotide exchange factor known as Sos protein, because of its homology with the exchange factor of *Drosophila* (Sos: son of sevenless, because of the role of this protein in signal transduction of the *sevenless* gene in *Drosophila*).

In addition to Sos protein, there are also other Ras-specific GEFs in mammals (see Table 9.1).

9.6.2 Structure and Activation of GEFs

Within each subfamily of the Ras superfamily, the GEFs are conserved and contain a distinct region to which nucleotide exchange activity can be assigned. The structure of the Ras-specific mSos protein (m=mammalian) is shown in Fig. 9.8. Nucleotide ex-

change activity is located within a domain of ca. 200 amino acids termed the *Cdc25 domain* by its homology to the Cdc25 exchange protein from the yeast *Saccharomyces cerevisiae*. The Cdc25 domain is characteristic of the Ras subfamily-specific GEFs. Other structural elements of mSos include a PH domain and a Pro-rich-binding domain. The Pro-rich sequence functions as an attachment site for the SH3 group of Grb2 protein.

The complex of Grb2 and mSos proteins forms a link between Ras protein and activated receptor tyrosine kinases or other transmembrane receptors. The Grb2 protein has two SH2 domains (see Section 8.5) and an SH3 domain.

The Sos-Grb2 complex can participate in Ras signal transduction by two pathways. In one pathway, the SH2 domain of Grb2 binds to the phosphotyrosine of the activated receptor, whereby the Grb2-mSos complex, which is predominantly localized in the cytoplasm, is brought to the receptor and thus to the cell membrane (Fig. 9.9).

In the other pathway, an additional adaptor protein, the Shc protein (see Section 8.5), is involved in the signal transduction. The Shc protein has a phosphotyrosine-binding domain (PTB domain) and specifically binds via this domain to autophosphorylated receptors such as the PDGF receptor and the EGF receptor. The Shc protein is phosphorylated itself in the process. The phosphotyrosine residues may then serve as attachment points for the SH2 domain of Grb2 protein, whereby the Grb2-Sos complex is attached to the membrane.

In the membrane-localized form, Sos protein interacts with Ras protein, which is also membrane associated, and induces nucleotide exchange in the latter. It is assumed that relocation of the Grb2-Sos complex from the cytosol to the membrane is the decisive step that establishes binding between the activated receptor and Ras protein. The membrane association of both proteins is sufficient for activation of signal transduction and to "switch on" the Ras protein, according to this assumption.

The exact mechanism of nucleotide exchange has been explained in broad outline for several GEF families revealing a similar mechanism of nucleotide exchange.

The Ras-specific GEFs bind to the GDP form of the Ras protein, thereby displacing GDP. The binary Ras-GEF complex may be dissociated by GDP or GTP. Since there is a large excess of GTP compared to GDP in the cell, the nucleotide-free Ras protein is preferentially converted to the GTP-bound, active state. The main driving force for nucleotide exchange is thus the high GTP concentration in the cell. By stabilizing

Fig. 9.8 Domain structure of mSos. The mSos-1 protein of mammals possesses a pleckstrin homology domain (PH), a Pro-rich domain for interaction with Grb2 and a catalytic domain with three sequence motifs (1,2,3) characteristic for Ras GEFs.

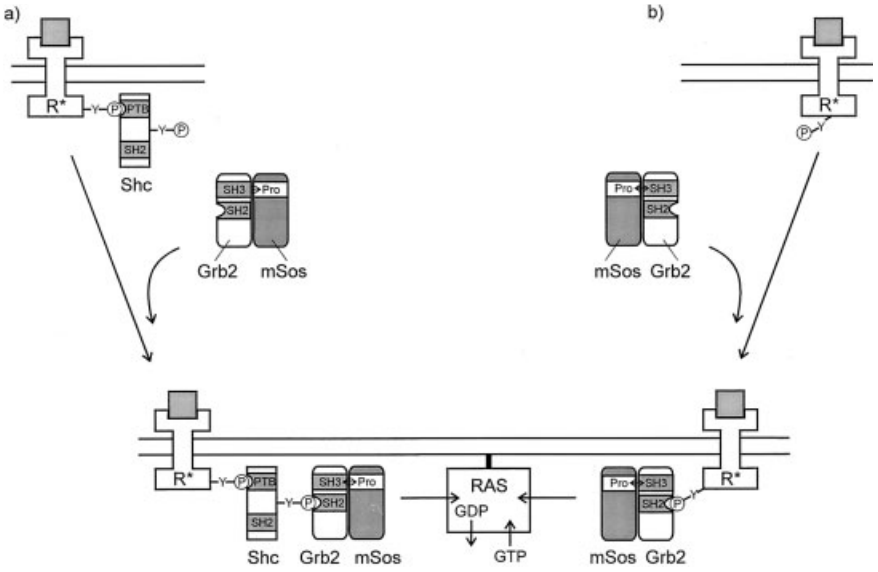


Fig. 9.9 Model of the function of Grb2-mSos and the adaptor protein Shc in the Ras pathway. The figure shows a highly simplified version of the two known pathways of involvement of the Grb2-mSos complex in signal transduction via the Ras protein. Phosphotyrosine residues of an activated, autophosphorylated receptor R* may serve as attachment points for the PTB domains of the Shc adaptor protein (a) or for the SH2 domain of the Grb2-mSos complex (b). In case (a), Tyr phosphorylation of Shc is performed by the activated receptor. The Grb2-mSos complex binds to the newly created phosphotyrosine residues and is drawn into the signal pathway. In case (b), the Grb2-Sos complex acts directly between the receptor and the Ras protein. In both situations, the Grb2-mSos complex is targeted to the membrane and from there, it can catalyze nucleotide exchange at the Ras protein.

the nucleotide-free state of Ras protein, GEFs enable binding of the G-nucleotide occurring at the highest concentration in the cell, namely GTP.

The mechanistic basis of nucleotide exchange by GEFs can be inferred from high-resolution structures between GEFs and members of the Ras-GTPase superfamily. The data show that GEFs interact mainly with the switch I and switch II regions. The P-loop and Mg^{2+} are displaced from binding to the phosphates by inserting GEF-residues into the nucleotide binding site of the Ras protein so as to sterically and electrostatically expel the nucleotide by a push-pull mechanims (see Fig. 9.10). The GEFs engage the switch II into an interaction and cause the displacement of switch I to open up the nucleotide binding site (Goldberg, 1998). The structural elements of Ras protein recognized by GEFs are identical in part to the sections to which other regulatory proteins bind, such as GAP protein and Raf kinase.

The GEFs for the Rho-family of small GTPases contain a domain responsible for nucleotide exchange that is different from the corresponding motif of the Ras-specific GEFs. This domain is called the Dbl homology (DH) domain. It has been named after the oncoprotein Dbl, which contains a domain of approximately 180 amino acids for which homologs were later found in a growing family of oncogenes. Proteins contain-

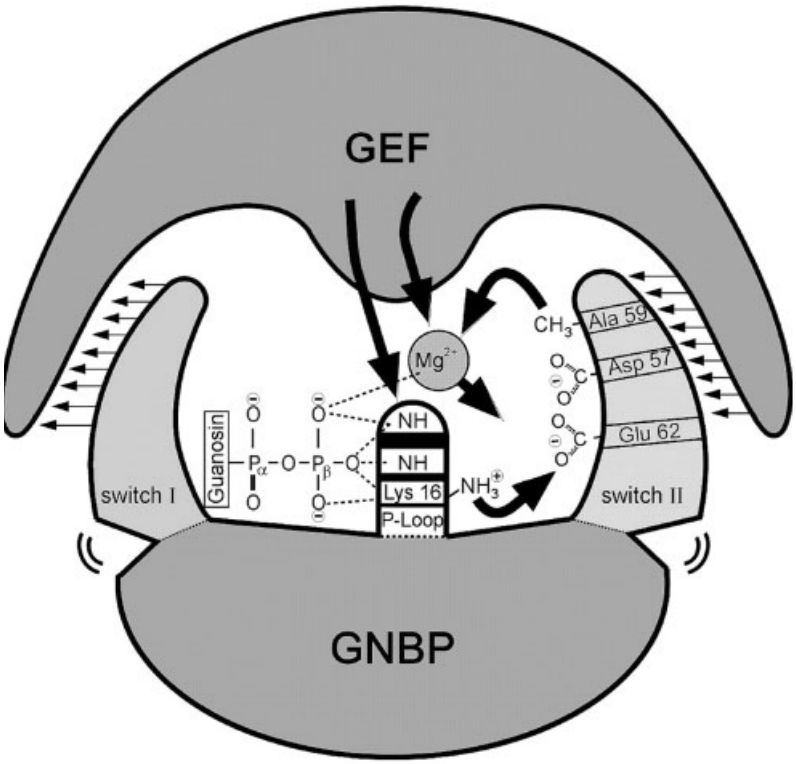


Fig. 9.10 Schematic diagram of the GEF action, showing common mechanistic principles. The most important contribution to high-affinity binding of the guanine nucleotide is due to interaction of the phosphates with the P loop and the Mg^{2+} ion. Mg^{2+} is pushed out of its position by elements of the GNBP itself, i. e., the Ala59 in Ras, or from residues of GEF. Residues of the P loop are disturbed, and its lysine is reoriented toward invariant carboxylates from the switch II region, either the invariant Asp57 in Ras or the highly conserved Glu62. In what might be called a push-and-pull mechanism, switch I is pushed out of its normal position, whereas switch II is pulled toward the nucleotide-binding site. From Vetter and Wittinghofer (2001). GNBP: guanine nucleotide binding protein.

ing the DH domain are now included in the Dbl protein family, which comprises >20 members (review: Whitehead et al., 1997). Important Dbl proteins are the Vav oncoprotein and the Bcr protein (see Section 14.3.3).

Regulation of the GEFs is very diverse and does not only include adaptor-mediated interactions with the activated receptor. There are other Ras-GEFs that are controlled by distinctly different mechanisms, namely via second messengers. Ras-GEFs have been identified in brain that are subject to control by diacylglycerol and Ca^{2+} (see Section 9.8). In addition, an exchange factor activated by cAMP has been described for Rap1 protein (see also Section 6.2; de Rooij et al., 2000). These examples show that the input signals to the Ras switching station can be very diverse.

9.7
Raf Kinase as an Effector of Signal Transduction by Ras Proteins

Which signal proteins are next in sequence after the Ras protein? The first effector protein identified for Ras was the Ser/Thr-specific protein kinase Raf. Following discovery of the viral oncogene *v-raf*, it was shown that cellular counterparts of it, the *c-raf* genes, code for a family of Ser/Thr-specific protein kinases that are activated by the GTP-form of the Ras protein and transmit a growth-promoting signal via the MAP kinase pathway (see Chapter 10) down to the transcriptional level. Historically, the identification of the Raf kinase as an effector of the Ras protein provided the first linkage between growth factor receptors, Ras signaling, and transcriptional activation during stimulation of cell proliferation by growth factors. By this discovery, it was possible for the first time to delineate a linear, apparently simply structured signal transduction pathway from the cell membrane down to the transcriptional regulation. However, it was recognized later on that other effector proteins are also used during Ras signaling providing linkages to other signaling pathways and that the Ras protein is part of a complicated signaling network (see Section 9.8).

The oncogenic *v-raf genes* differ from their cellular counterpart, the *c-raf* gene, mainly by deletions at the N terminus of the coded protein, affecting 200 – 300 amino acids. This section has an autoregulatory function for protein kinase activity of Raf protein. Loss of the autoregulatory domain is responsible for the transforming property of the viral protein. Mutations in the *raf* genes, mainly in the *B-raf* gene, are found in many tumors. As an example, about two thirds of malignant melanomas harbor a mutation in the *B-raf* gene. These mutations affect predominantly the catalytic domain of the B-Raf kinase, leading to enhanced kinase activity (review: Kolch, 2001).

9.7.1
Structure of Raf Kinase

Mammals have at least three different genes for Raf kinases, namely the genes A-*raf*, B-*raf* and c-*raf*1. The structure of c-Raf1 kinase is shown in Fig. 9.11. The three Raf kinases have three common conserved domains. Two of the domains, CR1 and CR2, are toward the N terminus and have a regulatory function on Raf activity. Specific mutations in these regions activate the transforming potential of Raf kinase. Particular importance for the Raf function is attributed to the CR1 domain. The CR1 domain has a Cys-rich section with the ability to bind Zn and phospholipids like phosphatidylserine and phosphatidic acid. These properties are reminiscent of the cysteine-rich domains found in the N-terminal region of protein kinase C providing a binding site for phorbol esters and for diacylglycerol (see Section 7.4). Mutations and deletions of the N-terminal sequences may bring about constitutive activation of Raf kinase. The CR1 element is thus considered to have an *autoregulatory function of Raf kinase activity*, and loss of this function may lead to oncogenic activation of Raf kinase. The CR2 domain contains Ser and Thr residues that serve as regulatory phosphorylation sites. Mutations are also described for these regions, leading to oncogenic activation of Raf kinase. The protein kinase activity is found in the CR3 domain.

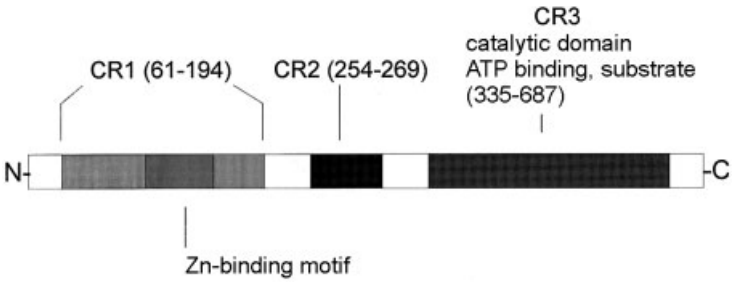


Fig. 9.11 Domain structure of Raf kinase. Linear representation of the functional domains of c-Raf1 kinase. CR: conserved region.

9.7.2
Interaction of Raf Kinase with Ras Protein

The activated GTP form of Ras protein interacts in a specific manner with Raf kinase and thus mediates membrane localization of Raf kinase. Consequently, the protein kinase activity of Raf kinase is stimulated and the signal is transmitted further via the protein kinase cascade of the MAP kinase pathway.

Signal transduction between Ras protein and Raf kinase is based on a specific interaction of the two proteins, which can only be performed by the activated, GTP-bound form of Ras protein.

Two structural elements of the Raf kinase have been found to interact with the Ras protein. One structural element involved in complex formation on the c-Raf1 kinase side is the Ras-binding domain, encompassing amino acids 51–131 of the CR1 domain. On the Ras protein side, the switch I region is involved in this interaction, as shown by the crystal structure of the complex between the Ras-binding domain of c-Raf1 kinase and the GTP form of the Ras-like protein, Rap1A protein (Nassar et al., 1995;). These structural studies did not reveal gross conformational changes of the two proteins upon complex formation.

Another region reported to bind to Ras protein encompasses amino acids 139–184, including the cysteine-rich region of CR1. It is still unknown which part of the Ras protein is involved in this contact and whether this interaction is required for Ras activation.

9.7.3
Mechanism of Activation and Regulation of Raf Kinase

Our understanding of the mechanism of activation of Raf kinase by Ras protein is still incomplete, mainly because full-length Raf kinase cannot yet be studied biochemically and structurally. Preliminary information was obtained in experiments using a fusion protein consisting of the membrane localization signal CAAX of K-Ras protein linked to the C terminus of Raf kinase. The presence of the membrane localization sequence of the Ras protein in Raf kinase leads to its constitutive activation, and the activity of

Raf kinase in this construct is independent of the Ras protein. In the K-Ras protein, the CAAX sequence represents a signal for farnesylation, which, together with a basic sequence (see Fig. 3.13), is sufficient to bring about membrane localization of the K-Ras protein. The experiment suggests that an important function of the activated Ras protein is to transport Raf kinase to the membrane in a regulated fashion (Fig. 9.12). It is still an open question how membrane translocation of Raf kinase is linked to its activation (review: Kerkhoff and Rapp, 2001). Possibly the binding of phospholipids to the cysteine-rich motif of the CR1 region of the Raf kinase is also involved in the activation step. It is assumed that the Ras-mediated membrane localization of Raf kinase is the first step which must be followed by other events, such as protein phosphorylation, oligomerization, and interaction with other cofactors, to bring about complete activation.

The complexity of regulation of Raf kinase was shown by the discovery of further proteins which can specifically associate with Raf kinase (reviews: Kerkhoff and Rapp, 2001; Chong et al., 2003). Members of the family of 14-3-3 proteins are found associated with Raf kinase. The 14-3-3 proteins recognize and bind phosphoserine residues in a particular sequence environment. It is assumed that the 14-3-3 proteins bind to Ser-phosphate residues of Raf kinase and thereby fix it in an inactive conformation. Other Raf-interacting proteins include the molecular chaperones hsp90 and p50. These proteins appear to be important for maintaining protein stability and for the proper localization of Raf kinase within the cell.

The activity of Raf kinase is also regulated by protein phosphorylation. Raf kinase is phosphorylated on Ser/Thr and Tyr residues. A negative regulation of Raf kinase is observed by protein kinase A-mediated phosphorylation. In addition, protein kinase C activates Raf kinase by phosphorylation. Furthermore, phosphorylation of Tyr residues of Raf kinase is seen in the process of activation of the Ras-Raf pathway. Src kinase is involved in the activating Tyr phosphorylation. How the different phosphorylation events are arranged in the overall pathway of Ras/Raf signal transduction is not well understood.

The path of signal transduction leads from the activated Raf kinase to the protein kinase cascade of the MAP kinase pathway (Chapter 10).

9.8 Reception and Transmission of Multiple Signals by Ras Protein

The Ras protein is a *multifunctional* signal protein that can be activated by various signaling pathways and transmits signals via different effector proteins. In addition to the well-characterized signaling pathways described above leading to the Raf kinase via growth factor receptors, other signal chains have been found that use Ras protein (or the different Ras subtypes) as a central switching station (see Fig. 9.13). These signaling pathways often show tissue specificity, and the details are not well characterized. The original simple picture of Ras signal transmission must therefore be replaced by a more complex picture in which the Ras protein receives many signals and transmits these to different effector proteins (reviews: Shields et al., 2000; Matozaki et

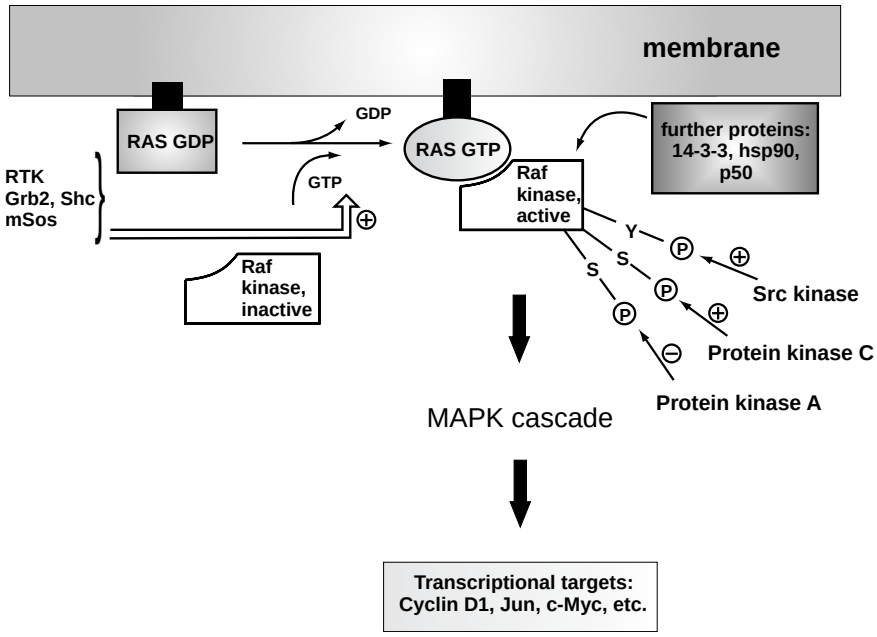


Fig. 9.12 Model of regulation and activation of Raf kinase. The active Ras.GTP complex binds and activates Raf kinase, which passes the signal on to the MAP kinase pathway. Various proteins including the 14-3-3 proteins and the molecular chaperons hsp 90 and p50 are thought to be involved in the regulation of the Raf kinase signaling function. In addition, Raf kinase is regulated by phosphorylation. Tyr phosphorylation (possibly via Src kinase) and Ser phosphorylation via protein kinase C have a stimulatory effect. In contrast, Ser phosphorylation via protein kinase A has an inhibitory effect. RTK: receptor tyrosine kinase.

al., 2000). It is not known how the various activities in this Ras network are coordinated. Furthermore. linkages between members of the Ras subfamily and members of the Rho/Rac subfamily have been discovered. Signals received by the Ras protein may thereby be passed on to the Rho/Rac proteins, allowing for, e. g., reorganization of the cytoskeleton in response to growth factor stimulation.

Multiple Input Signals of Ras Protein

The Ras protein may be activated by different signaling pathways:

- **Binding of growth factors to receptor tyrosine kinases**
  This well-characterized pathway of Ras signal transmission was the first to be discovered (see above) and involves adaptor proteins (Grb2, Shc) and GEFs (e. g., mSos).
- **Binding of cytokines to receptors with associated tyrosine kinase activity; activation of integrins**

- **Ca^{2+} and diacylglycerol signals**
  Changes in the concentration of diacylglycerol and Ca^{2+} lead to activation of the Ras protein in brain. This effect is possibly mediated via specific GEFs. Ras-specific GEFs, which are regulated by Ca^{2+}, are found in brain. Examples are the Ras guanyl nucleotide-releasing protein (RasGRP), which contains a Ca^{2+}-binding motif, a diacylglycerol-binding motif, and the Ras guanyl nucleotide-releasing factor 1 (RasGRF1), which is activated by Ca^{2+}/calmodulin (Ebinu et al., 1998).
- **NO signals**
  Stimulation of *N*-methyl-D-aspartate (NMDA) receptors in the nervous system is linked to activation of NO synthase and creation of an intracellular NO signal (Yun et al., 1998). NO can directly activate Ras protein by redox-modification; S-nitrosylation (see Section 6.10) of Cys 118 of Ras has been shown to trigger GDP/GTP exchange and to convert Ras into the active GTP state.
- **Signals from G-protein signaling pathways**
  There is evidence that signals starting from G-protein-coupled receptors run into the Ras switch station. $\beta\gamma$-subunits of G-proteins are under discussion as the link between G-protein-coupled signal transduction and the Ras pathway; these subunits could influence the activity of Ras protein and the subsequent MAP kinase pathway by a presently unknown mechanism.

Multiple Effector Molecules of Ras Proteins

Operationally, Ras effector proteins are characterized by their preferential binding to the active GTP form of Ras as compared to the inactive GDP form. In addition to Raf kinase, a number of other signal proteins have been identified to which an effector function in Ras signal conduction has been attributed (review: Shields et al., 2000). These effector candidates include a very diverse collection of structurally and functionally distinct proteins, which all show preferential affinity for the active Ras form. The residues of Ras protein involved in binding these effectors lie in the region of switch I (residues 30–37) and switch II (residues 59–76); in addition, an intact Ras effector domain (residues 32–40) is required for this interaction.

- **Raf kinase (see above)**
  Most of the biological effects are mediated by activation of Raf kinase, which is part of the MAPK/ERK signaling module (see Chapter 10).
- **MEK kinases**
  In addition to Raf kinase activation, Ras protein also mediates stimulation of other protein kinases known as MEK kinases. These are signal proteins in another MAPK module, the JNK signaling pathway (see Chapter 10), and transmit signals at the level of gene expression.
- **PI3-kinase**
  The GTP form of Ras protein specifically binds to the catalytic 110 kDa subunit of phosphatidyl inositol-3-kinase (PI3-kinase, see Section 6.6). There is clear experimental evidence that PI3-kinase is downstream from the Ras protein and that its activity may be controlled by Ras protein. Activation of PI3-kinase leads to

the formation of the membrane-localized messenger substance Ptd-Ins(3,4)P_3, which binds to the PH domains of signal proteins and can lead these to the membrane and activate them (see Section 6.6).

The Ras*GTP-mediated activation of PI3-kinase links the Ras pathway with functions of the Rho/Rac proteins. Members of this protein family within the Ras super-

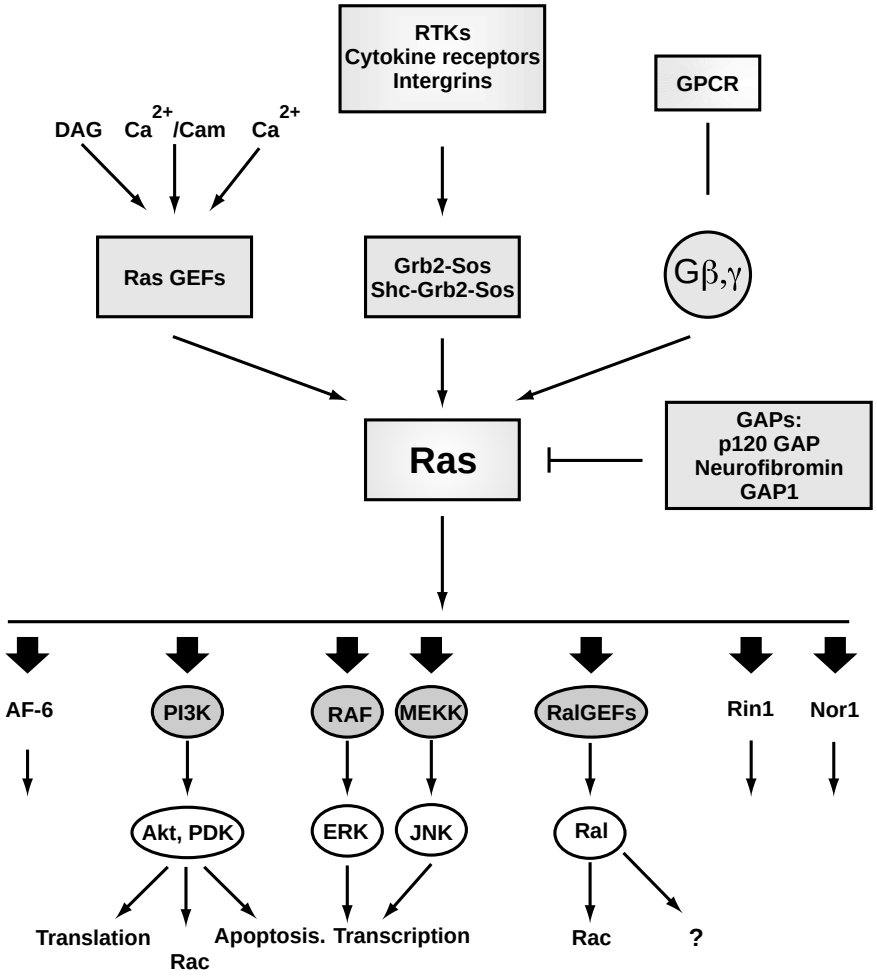


Fig. 9.13 Summary of input and ouput signals of Ras. Input signals originate mostly from Ras-GEFs, Shc-Grb2-mSos complexes and from $\beta\gamma$ complexes. A negative regulation of Ras occurrs by various GAPs. Activated Ras delivers signals to the level of transcription via MAPK cascades and thereby activates an array of genes of which many promote cell proliferation. Another important proliferation promoting and antiapoptotic signal can be delivered from Ras via the PI3/Akt kinase pathway. Linkages of Ras to the Rac proteins are mediated by Ral-GEFs and by the Akt kinase pathway. The pathways activated by AF-6, Rin1 and Nor1 are largely unknown.

family control formation of the cytoskeleton. The exact nature of the linkage with the Ras/PI3-kinase signal conduction to the Rac proteins is unknown. There is evidence that the product of the PI3-kinase, Ptd-Ins(3,4,5)P_3, binds to the PH domain of the Vav protein and activates the latter. The Vav protein functions as a nucleotide exchange factor for the Rac GTPase (review: Bustelo, 2001). The observation that activation of the Ras pathway is accompanied by reorganization of the cytoskeleton is in agreement with these findings.

Ras-mediated activation of PI3-kinase also links the Ras protein to Akt kinase (see Section 6.6.1), which mediates antiapoptotic signals.

- **GEFs of Ral GTPase**

  Proteins that function as GEFs for the Ral GTPase have also been identified as effectors of Ras proteins. The Ral protein is a Ras-related small GTPase that is involved in, e. g., activation of phospholipase D1 and inhibition of the Rac proteins. The GEFs are proteins known as Ral-GDS, RLF and Rgl (review: Wolthuis and Bos., 1999).

  The Ral-specific GEFs are activated by the interaction with Ras protein. It is assumed that activation is primarily caused by membrane association coupled to Ras binding.

Other less well-characterized effectors of Ras proteins are the proteins Rin1, Nore 1 and AF6 (review: Vojtek and Der, 1998). Furthermore an effector function has also been described for the Ras-activating protein p120 GAP, which, in addition to negatively regulating the Ras function, also specifically associates with the protein p190, which is a GAP for the Rho family of GTPases. It is assumed that activation of the Rho family of GTPase contributes significantly to the Ras-transformed phenotype.

Ras Protein as a Central Switching Station

Identification of multiple input signals and several effector proteins underlines the high complexity of signal transduction via the Ras protein. The Ras pathway cannot be seen as a linear ordering of signal elements by which information is conducted vertically from the cell membrane to the cell interior. Rather, the Ras protein is at the center of a network of different signal chains (see Fig. 9.13). In this network, it functions as a central switching station at which signals are registered, integrated and passed on.

Different signals meet at the Ras switching station. Signals starting from activated receptor and nonreceptor tyrosine kinases are registered. These signals may be conducted via GEFs and/or GAP proteins to the Ras protein. Ca^{2+} signals, redox signals in the form of NO, and signals from G-protein-signaling pathways are also received at the Ras switching station.

Starting from the activated Ras protein, compounds are produced for various signaling pathways. A main pathway leads via MAP kinase to the level of gene expression, creating proliferation-promoting or also -inhibiting signals. It is assumed that the transforming effect of oncogenic Ras mutants is mediated by this pathway in particular. A function-promoting cell survival is also mediated via linkage to the PI3-Akt pathway.

Important issues of Ras functions remain to be clarified. The spectrum of known Ras-activated genes is certainly still very incomplete, and the tissue-specificity and cell-specificity of Ras proteins are ill defined. Furthermore, distinct members of the Ras subfamily appear to perform distinct functions that may overlap with other Ras proteins. Even within the genuine Ras proteins, the H-Ras, K-Ras and N-ras proteins, it has not yet been possible to discover functional differences.

Surprisingly, a growth-inhibiting and pro-apoptotic function has been demonstrated for oncogenic Ras mutants. In primary cell cultures, activation of the Ras pathway is linked to an increase in the concentration of the tumor suppressor proteins p53 and p19ARF (Serrano, 1997), which both promote programmed cell death or apoptosis (see Chapter 15). It is now increasingly recognized that cells contain safeguard mechanisms that can couple the overstimulation of growth-promoting pathways to the induction of apoptosis (see Chapter 15).

There are also links to other members of the Ras superfamily such as the Ral protein and the Rho/Rac proteins. The latter are involved in reorganization of the actin cytoskeleton. Transformation of cells with oncogenic Ras mutants is associated with reorganization of the actin cytoskeleton, and it is assumed that this effect is due to coupling of the Ras pathways with the function of Rho/Rac proteins.

Reference

1 Boriack-Sjodin, P.A., Margarit, S.M., Bar-Sagi, D. and Kuriyan, J. (1998) The structural basis of the activation of Ras by Sos. *Nature*, **394**, 337–343.

2 Brose, M.S., Volpe, P., Feldman, M., Kumar, M., Rishi, I., Gerrero, R., Einhorn, E., Herlyn, M., Minna, J., Nicholson, A. *et al.* (2002) BRAF and RAS mutations in human lung cancer and melanoma. *Cancer Res.*, **62**, 6997–7000.

3 Bustelo, X.R. (2001) Vav proteins, adaptors and cell signaling. *Oncogene*, **20**, 6372–6381.

4 Chong, H., Vikis, H.G. and Guan, K.L. (2003) Mechanisms of regulating the Raf kinase family. *Cell Signal.*, **15**, 463–469.

5 de Rooij, J., Rehmann, H., van Triest, M., Cool, R.H., Wittinghofer, A. and Bos, J.L. (2000) Mechanism of regulation of the Epac family of cAMP-dependent RapGEFs. *J.Biol.Chem.*, **275**, 20829–20836.

6 Ebinu, J.O., Bottorff, D.A., Chan, E.Y., Stang, S.L., Dunn, R.J. and Stone, J.C. (1998) RasGRP, a Ras guanyl nucleotide- releasing protein with calcium- and diacylglycerol-binding motifs. *Science*, **280**, 1082–1086.

7 Ellis, C.A. and Clark, G. (2000) The importance of being K-Ras. *Cell Signal.*, **12**, 425–434.

8 Garbay, C., Liu, W.Q., Vidal, M. and Roques, B.P. (2000) Inhibitors of Ras signal transduction as antitumor agents. *Biochem.Pharmacol.*, **60**, 1165–1169.

9 Geyer, M. and Wittinghofer, A. (1997) GEFs, GAPs, GDIs and effectors: taking a closer (3D) look at the regulation of Ras-related GTP-binding proteins. *Curr.Opin.Struct.Biol.*, **7**, 786–792.

10 Goldberg, J. (1998) Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell*, **95**, 237–248.

11 Kerkhoff, E. and Rapp, U.R. (2001) The Ras-Raf relationship: an unfinished puzzle. *Adv.Enzyme Regul.*, **41**, 261–267.

12 Kolch, W. (2001) To be or not to be: a question of B-Raf? *Trends Neurosci.*, **24**, 498–500.

13 Kuersten, S., Ohno, M. and Mattaj, I.W. (2001) Nucleocytoplasmic transport: Ran, beta and beyond. *Trends Cell Biol.*, **11**, 497–503.

14 Nassar, N., Horn, G., Herrmann, C., Scherer, A., McCormick, F. and Wittinghofer, A. (1995) The 2.2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. *Nature*, **375**, 554–560.

15 Nassar, N., Horn, G., Herrmann, C., Scherer, A., McCormick, F. and Wittinghofer, A. (1995) The 2.2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. *Nature*, **375**, 554–560.

16 Price, L.S. and Collard, J.G. (2001) Regulation of the cytoskeleton by Rho-family GTPases: implications for tumour cell invasion. *Semin.Cancer Biol.*, **11**, 167–173.

17 Reuther, G.W. and Der, C.J. (2000) The Ras branch of small GTPases: Ras family members don't fall far from the tree. *Curr.Opin.Cell Biol.*, **12**, 157–165.

18 Scheffzek, K., Ahmadian, M.R., Kabsch, W., Wiesmuller, L., Lautwein, A., Schmitz, F. and Wittinghofer, A. (1997) The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*, **277**, 333–338.

19 Segev, N. (2001) Ypt/rab gtpases: regulators of protein trafficking. *Sci.STKE.*, **2001**, RE11.

20 Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D. and Lowe, S.W. (1997) Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell*, **88**, 593–602.

21 Shields, J.M., Pruitt, K., McFall, A., Shaub, A. and Der, C.J. (2000) Understanding Ras: 'it ain't over 'til it's over'. *Trends Cell Biol.*, **10**, 147–154.

22 Takai, Y., Sasaki, T. and Matozaki, T. (2001) Small GTP-binding proteins. *Physiol Rev.*, **81**, 153–208.

23 Takai, Y., Sasaki, T. and Matozaki, T. (2001) Small GTP-binding proteins. *Physiol Rev.*, **81**, 153–208.

24 Vetter, I.R. and Wittinghofer, A. (2001) The guanine nucleotide-binding switch in three dimensions. *Science*, **294**, 1299–1304.

25 Vojtek, A.B. and Der, C.J. (1998) Increasing complexity of the Ras signaling pathway. *J.Biol.Chem.*, **273**, 19925–19928.

26 Whitehead, I.P., Campbell, S., Rossman, K.L. and Der, C.J. (1997) Dbl family proteins. *Biochim.Biophys.Acta*, **1332**, F1–23.

27 Wolthuis, R.M. and Bos, J.L. (1999) Ras caught in another affair: the exchange factors for Ral. *Curr.Opin.Genet.Dev.*, **9**, 112–117.

28 Yun, H.Y., Gonzalez-Zulueta, M., Dawson, V.L. and Dawson, T.M. (1998) Nitric oxide mediates N-methyl-D-aspartate receptor-induced activation of p21ras. *Proc.Natl.Acad.Sci.U.S.A*, **95**, 5773–5778.

10
Intracellular Signal Transduction: the Protein Cascades of the MAP Kinase Pathways

Intracellular signal conduction takes place predominantly by two pathways starting from activated transmembrane receptors. In one pathway, activation of transmembrane receptors initiates formation of diffusible messenger substances that bind effector proteins and activate these for further signal transduction. In this signaling pathway, signals may be carried as far as the cell nucleus, and temporally and spatially variable reactions may be triggered.

A second pathway, which is particularly important for regulation of growth and differentiation reactions, takes place via a *cascade* of sequential protein kinases. In this pathway, an extracellular signal is registered by a transmembrane receptor, which is activated and transmits the signal to downstream effector molecules. The Ras protein or other members of the Ras protein superfamily are often involved in this pathway. Up to this point, all the reactions involved are predominantly membrane associated. From the Ras protein (or other regulatory proteins), the signal is conducted, with the help of protein kinases, in the form of a sequential cascade, into the cell interior, possibly into the cell nucleus (Fig. 10.1). Since these signaling pathways are often activated by mitogenic (mitogenic = promoting cell division activities) signals such as growth hormone signals, they are known as mitogenic activated protein kinase pathways (MAPK pathways). As a result of stimulation of the MAPK pathways, phosphorylation and activation of gene-regulating proteins or enzymes catalyzing key reactions of metabolism are observed (review: Robinson and Cobb, 1997, Kyriakus and Avruch, 2001). Signal transduction via *sequential protein kinase reactions* is a very flexible and efficient principle for amplification, diversification, and regulation of signals. Protein kinases, as explained in Chapter 7, are open to a range of regulatory influences. At every level of a protein kinase cascade, positive or negative regulation is possible, and the intensity of a signal can be modulated within broad boundaries.

Organization of MAPK Pathways in MAPK Modules

The mitogenic activated protein kinase (MAPK) pathways are organized in modules containing at least three types of protein kinases, which transmit signals by sequential phosphorylation events in a hierarchical way. A multitude of signals have been found to activate MAPK modules. Historically, the activation of the Raf kinase via Ras proteins was discovered first as an entry point for the MAPK cascade signaling, providing a path from activated growth factors via Ras and Raf kinase down to the transcriptional

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

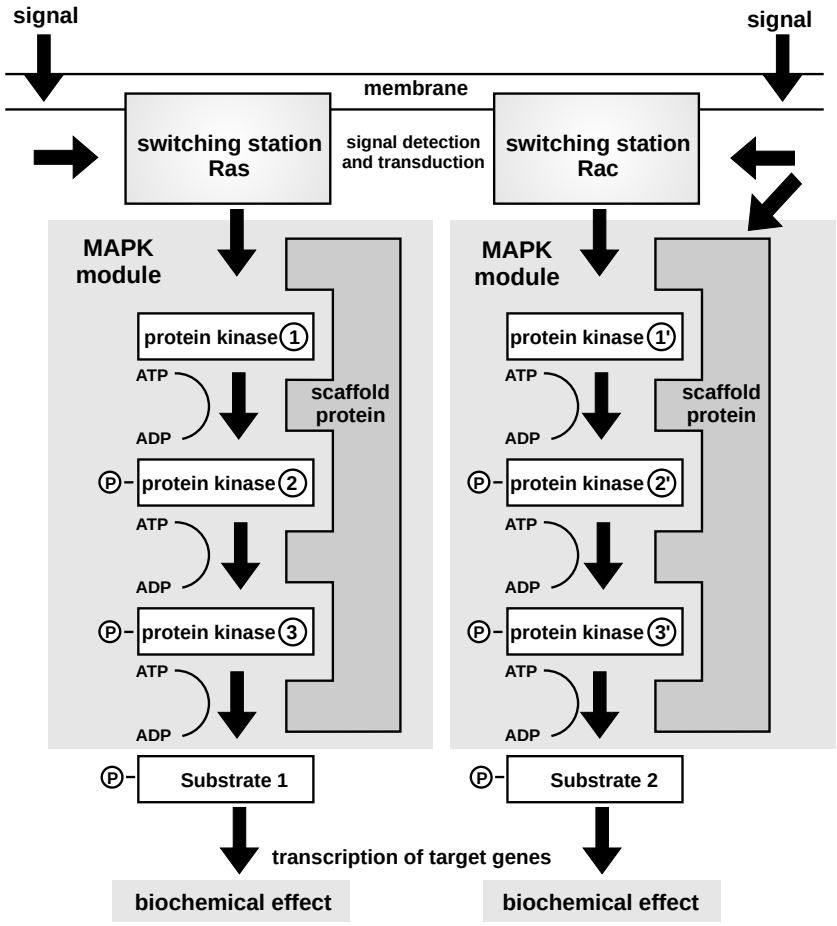


Fig. 10.1 Principle of signal transduction through intracellular protein kinase cascades. The intracellular protein kinase cascades are organized in modules composed in most cases of three proteinkinases and a scaffolding protein. The modules process signals that are registered, integrated and passed on at the inner side of the cell membrane by central switching stations such as the Ras protein or the Rac protein. In the case of the MAP kinase pathway, the cascade includes at least three different protein kinases. Specific regulatory processes may take effect at every level of the cascade; in addition, signals may be passed from the different protein kinases to other signaling pathways.

level. Later on it became clear that the cell contains different MAPK modules, which differ in the nature of the triggering stimuli and the nature and specificity of the protein kinase components. The signal-transducing function of a MAPK pathway is thus determined by the nature of the MAPK module involved; this, in turn, depends on the properties of the protein kinases it contains, which differ in regulation and substrate specificity. The exact composition of the MAPK module is not fixed; rather, different subtypes of protein kinase may be recruited to a module in a variable manner. Furthermore, the modules can share a common protein kinase but regulate different biolo-

gical processes. The signal is passed on by the last member in the phosphorylation cascade in the form of a phosphorylation of substrate proteins. In many cases, this process is linked to translocation of the protein kinase into the nucleus, where nuclear localized substrates, particularly transcription factors, are phosphorylated.

The various modules are not strictly independent; rather they mutually influence one another. In mammals in particular, this results in high complexity of function and regulation of MAPK pathways.

10.1 Components of MAPK Pathways

MAPK pathways comprise a core module of three protein kinases that are organized as functional units with the aid of scaffold proteins (see below). Incoming signals are received by the top (or uppermost) protein kinase of the core module and are passed on to substrates by the terminal protein kinase. For signal transduction starting from Ras protein (or other members of the Ras superfamily), the protein kinases at the lower end of the cascade, because of their regulation by extracellular, frequently mitogenic, ligands, are known as *extracellular regulated kinases* (= *ERK*) and sometimes as *mitogen-activated protein kinase* (= *MAP kinase, MAPK*). The term MAPK is often collectively used for the terminal kinase of the MAPK cascade, referring to several different terminal protein kinases.

One of the first MAPK pathways to be characterized leads from mitogens, via the Ras protein, to activation of protein kinases known as ERK1 and ERK2. This pathway is known as the *ERK pathway*. It should be noted that there are different MAP kinases which are active in other MAPK pathways or MAPK modules; these are different from the ERK pathway (see below). In the following, the ERK pathway starting from the activated Ras protein is used to represent other MAPK pathways with similar structure.

The MAPK/ERK proteins are at the lower end of signal transduction within a MAPK module and are generally preceded by two other protein kinases (Fig. 10.2). The MAPK/ERK proteins receive the signal in the form of an activating phosphorylation by a preceding protein kinase known as *MAP/ERK kinase* (*MEK*) or also *MAP kinase kinase* (*MAPKK*).

The MEK proteins are themselves substrates for another type of protein kinase further upstream, the *MEK kinases* (*MEKKs*, also known as *MAPKK kinases*, *MAPKKKs or MAP3Ks*).

The position of the various protein kinases within a MAPK module is determined by their substrate specificity.

MAP3Ks, MEK Kinases

The MAP3Ks (MAPKK kinases or MEK kinases) are Ser/Thr-specific protein kinases and are the entry point for signal transduction in a MAPK module. The best-characterized representative, Raf-1 kinase, is activated by Ras protein in its GTP-bound form. Raf kinase phosphorylates downstream MEK proteins at two Ser residues, which are separated by three other amino acids. All known MEK proteins have a similar phosphorylation site in

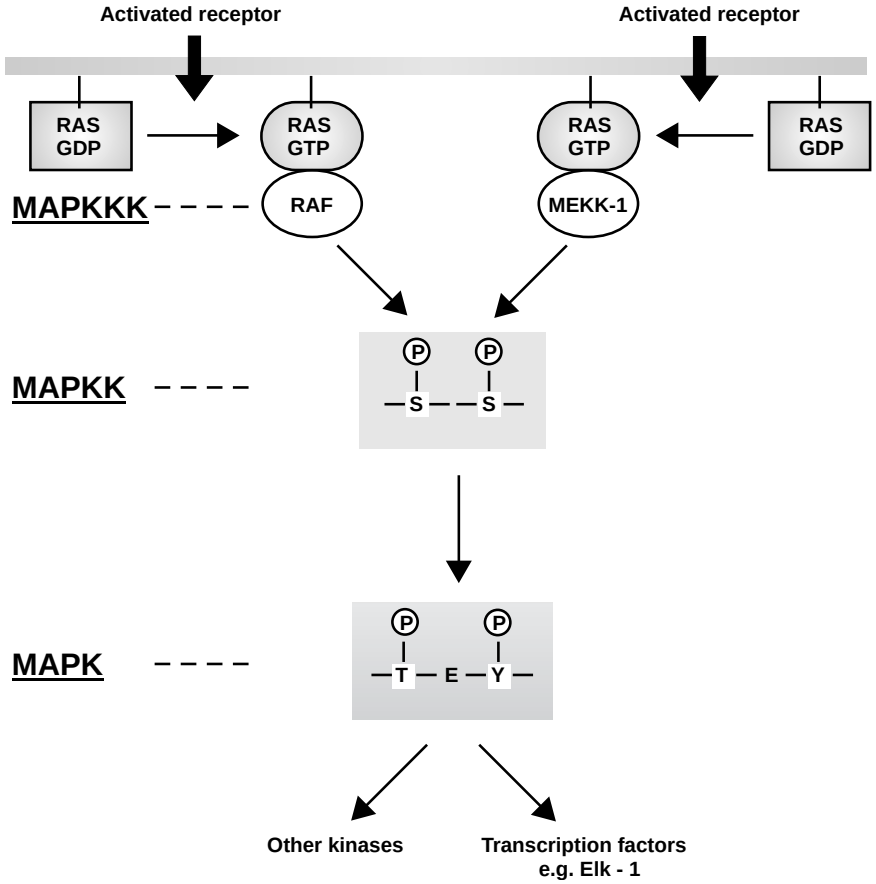


Fig. 10.2 Components and activation of the ERK pathway. Ordering and specificity of protein kinases in the ERK pathway. Extracellular signals are registered via receptor tyrosine kinases and passed on to the Ras protein. Ras-GTP activates protein kinases belonging to the group of MAPKK kinases (Raf kinases and MEKKs). The MAPKK kinases phosphorylate the downstream group of protein kinases, the MAPKKs at two Ser residues. The MAPKKs phosphorylate the MAPKs (ERK1 and ERK2) at a Tyr and a Thr residue, and thusare classified as daul specificity protein kinases.. MAPK: mitogenic activated protein kinase; ERK: extracellularly regulated kinase; MEK: MAP/ERK kinase; MAPKK: MAPK kinase; MAPKKK: MAPKK kinase; MEKK: MEK kinase.

the conserved sequence LID/NSXANS/T (X: any amino acid). Other representatives of the MEK kinase group are Mos kinase and the protein kinases MEKK1-3.

Signaling proteins that deliver signals to the MAP3Ks comprise mostly small GTPases like members of the Ras and Rho/Rac family or other protein kinases. In the latter case the MAP3Ks are activated by upstream protein kinases, which are then also referred to as MAP4Ks. In general, activation of MAP3Ks is a complex process requiring the steps of membrane translocation, phosphorylation, oligomerization, and binding to scaffold proteins. Mechanistic details of MAP3K activation are still not understood.

MAPK Kinases, MEKs

The MEK proteins are a special class of protein kinases, since they have twofold specificity with respect to the nature of the acceptor amino acid at the phosphorylation site of the protein substrate (review: Hagemann and Blank, 2001). The MEKs activate the MAPK/ERK proteins next in sequence by phosphorylation at a Tyr and a Thr residue in the TXY sequence. The MEKs are *dual-specificity protein kinases* and, because of this property, differ significantly from the other protein kinases described so far, which are either Tyr or Ser/Thr-specific.

MAP Kinases, ERKs

The MAP kinases are divided, based on their sequence homology, the nature of the preceding MEK, and the input signal, into different subgroups characteristic of the particular MAPK module. Within these subgroups, further diversification is possible by alternative splicing. Most of the known MAP kinases contain a TXY sequence; phosphorylation of this by MEKs is essential for conduction of the signal further to downstream substrate proteins. The TXY motif is found as part of the activation loop, and its phosphorylation by MEKs activates the MAPK enzymes by relieving a steric hindrance to substrate binding and by reorganization of the catalytic site.

Location of the substrates of the MAPKs can be nuclear or cytoplasmic. Recognition and selection of substrate proteins occurs via specific docking sites on the substrates, which are bound by complementary binding domains on the MAPK (review: Barsyte-Lovejoy et al., 2002). These docking sites are located at some distance from the phosphorylation site.

Organization in Multiprotein Complexes

There is increasing experimental evidence that coupling of the activities of the MAPK cascade is achieved with the help of organization in multiprotein complexes by employing specific scaffold proteins. Scaffolding of the components of the MAPK pathways has several advantages:

- It favors the rapid transmission of the signal through the cascade.
- It prevents unwanted crosstalk between different MAPK modules.
- It increases specificity and selectivity of signaling by the assembly of distinct kinases into distinct MAPK modules, thereby ensuring the coordinated and efficient activation of MAPK components in response to diverse stimuli.

There are several types of scaffold proteins. Some function simply as adaptors, while others perform additional functions. Furthermore, components of the MAPK modules can themselves perform the scaffolding function.

A scaffold and adaptor function could be ascribed to the Ste5 protein from the yeast *Saccharomyces cerevisiae*, which recruits three protein kinases of the pheromone-signaling pathway into a distinct MAPK module. The role of scaffolding in mammalian MAPK signaling is less well-characterized, although a number of candidate scaffolding proteins have been described in mammals (review: Kyriakis and Avruch, 2001; Chona et al., 2003). Examples are the MP1 protein that binds specifically to ERK1 and MEK1, the c-Jun-N-terminal kinase interacting proteins (JIP1s) and the SAPK (see below)

associated proteins (JSAP1/JIP3). Another candidate scaffold protein is arrestin (see Section 5.3.4), which organizes kinases of the SAPK group into distinct multiprotein complexes, providing for a link between G-protein-coupled receptors and the MAPK cascade (McDonald et al., 2000). Furthermore, a scaffolding function has been reported for a putative downstream effector of Ras, the kinase suppressor of RAS, KSR, which appears to organize Raf kinase into a signaling complex at the cell membrane (Raabe and Rapp, 2002).

10.2 The Major MAPK Pathways of Mammals

Cells contain several distinct MAPK-signaling pathways. In the yeast *S. cerevisiae*, six distinct MAPK signaling pathways have been identified to date. Of the at least six mammalian pathways, the ERK, JNK/SAP and the p38 pathways stand out and will be presented in more detail in the following discussion (Fig. 10.3).

10.2.1 The ERK Pathway

The ERK pathway was the first MAPK pathway to be identified. This pathway has been known for its activation by Ras proteins, which recruit MAP3Ks of the Raf family to activate two MEKs, MEK 1 and MEK 2. These in turn activate two ERKs, ERK 1 and ERK 2.

Input Signals

The following stimuli have been shown to activate and regulate the ERK pathway; some of these are active only in specific cells, while others operate in most cell types:

- Mitogenic signals originating from growth factor receptors (see Chapter 8); via Ras
- Mitogenic signals originating from cytokine receptors (see Section 11.2.1); via Ras or Src kinase, which has been shown to activate Raf kinase.
- Signals from integrins (see Chapter 11); via Ras and Src kinase
- Signals from Rho/Rac proteins
- Activation of G-protein-coupled receptors. There are many routes by which ligand-binding to G-protein-coupled receptors can transmit signals to the ERK pathway (review: Liebmann, 2001), and activation of the ERK pathway is frequently observed upon ligand binding to G-protein-coupled receptors. A main entry point is the Raf kinase, which can be activated by protein kinase C and inhibited by protein kinase A. As outlined in Sections 7.3 and 7.4, both enzyme families can be activated via G-protein-signaling pathways by multiple mechanisms.

Most of the incoming signals of the ERK pathway are received by the Raf kinase. Other MEKKS of this pathway include the Mos kinase and the MEKK-1 protein, which is also found as a component of the JNK/Sap and the p38 pathway.

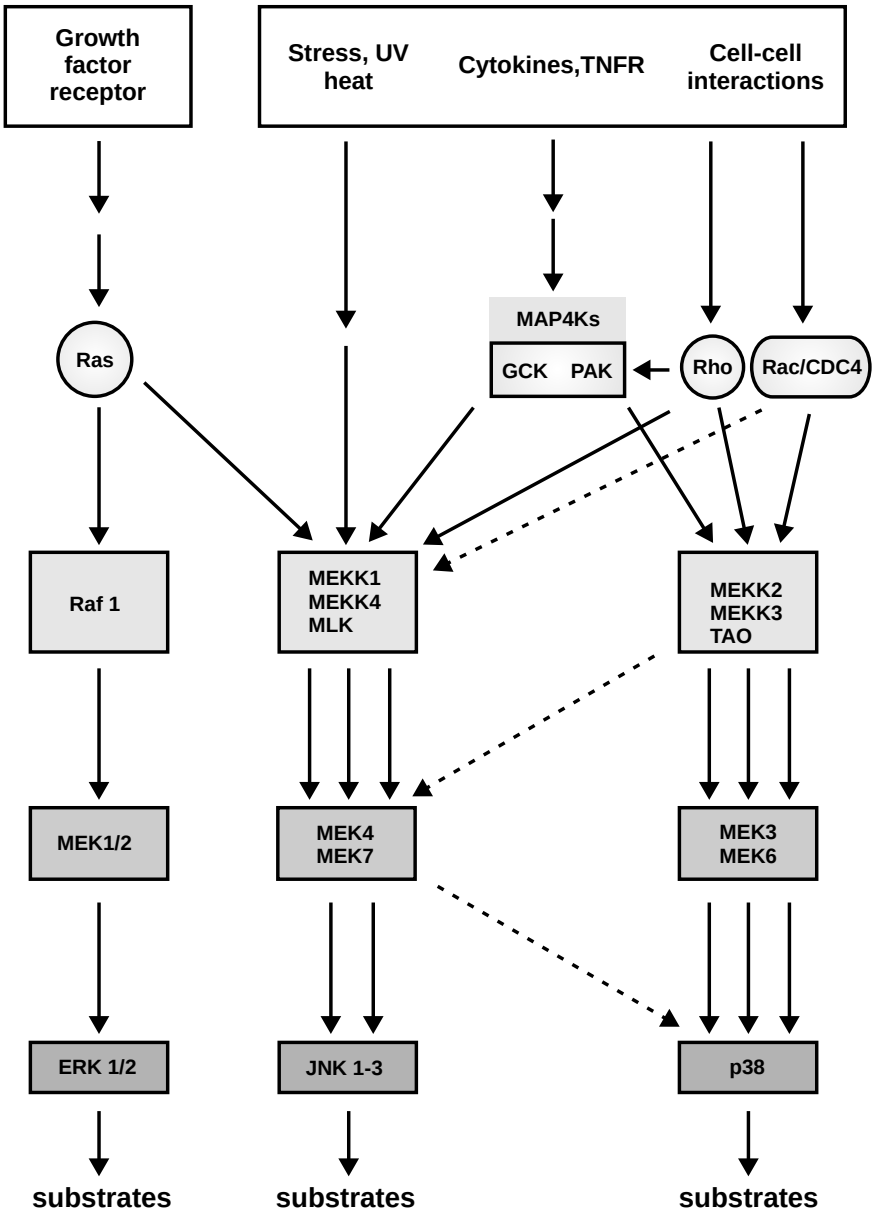


Fig. 10.3 Summary of the three major MAPK pathways in mammals. In addition to the Ras/Raf/MEK/ERK pathway, which are activated by receptor tyrosine kinases, there are two major MAPK pathways, the SAPK/JNK pathway and the p38 pathway, which are activated by other signals, e. g., UV irradiation, heat stress and cell-cell interactions. How these other signals are registered and passed to the small regulatory GTPases or to downstrem protein kinases is only incompletely known. Protein kinases of the MEKK (or MAP3K) class may be activated by small regulatory GTPases or by protein kinases classified as MAP4Ks. The figure illustrates the multiple interactions between the three pathways. Rac-1 and Cdc42 belong to the Rho family of small regulatory GTPases. TNFR: tumor necrosis factor receptor; GCK: germinal center kinase; PAK: p21-activated kinase; MLK: mixed lineage kinase; TAO: thousand and one-amino acid protein kinase.

Substrates of ERKs

ERKs deliver signals both to nuclear and cytoplasmic substrates. Activation of ERKs by MEK-mediated phosphorylation promotes their dimerization (review: Cobb and Goldsmith, 2000), and the ERK dimers can then translocate into the nucleus, where various transcription factors are phosphorylated and activated. An example of a nuclear substrate is the transcription factor Elk-1, which is positively regulated via the ERK pathway (review: Yordi and Muise-Helmericks, 2000). Elk-1 binds, together with another protein, the *serum response factor* to the *serum response element*, which is found as a regulatory sequence preceding various genes in higher eucaryotes. The genes regulated by Elk-1 include the gene for the transcription factor c-fos, which forms transcription factor AP1 together with the c-Jun protein (see Chapter 1). Elk-1 is phosphorylated by ERK proteins specifically at the sites essential for transcription activation. Several signals meet at the level of Elk-1, since activation of Elk-1 is mediated by different

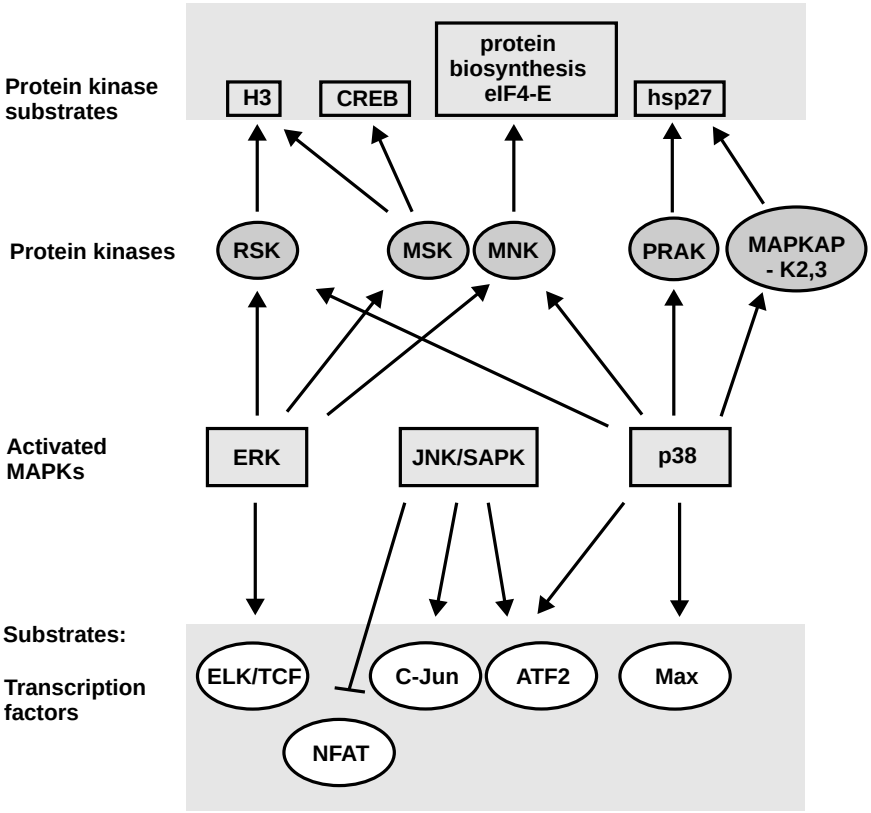


Fig. 10.4 Substrates of the major MAPK pathways. The major substrates of activated MAPKs are trasncription factors (bottom) and protein kinases that phosphorylate and regulate a multitude of substrates (top). MAPKAP-K: MAPK-activated protein kinase; MNK: mitogen-activated protein kinase-interacting kinase; MSK: mitogen- and stress-activated protein kinase; PRAK: p38-regulated/activated protein kinase; RSK: 90kDa-ribosomal S6 kinase.

MAPK proteins, which in turn are activated by different MAP kinase pathways (see Fig. 10.4). Much of the mitogenic effects of ERK signals can be explained by the observation that ERK-induced expression of transcription factors stimulates the transcription of D-type cyclins, which promotes the G1/S-phase transition in the cell cycle (see Section 13.4.1). Furthermore, phosphorylation of the CTD of RNA polymerase II (see Section 1.4.3.4) has also been reported to be mediated by the ERK1/2 proteins.

Phospholipase A2 is an example of cytoplasmic ERK substrates. Phosphorylation of a Ser residue of phospholipase A2 by ERK proteins leads to activation of the lipase activity. Consequently, there is an increase in release of arachidonic acid and of lysophospholipids, which can act immediately as diffusible signal molecules or may represent first stages in the formation of second messenger molecules.

An important cytoplasmic substrate of ERK1/2 proteins is the 90 kDa ribosomal S6 kinase (RSK), also termed MAPK-activated protein kinase 1, MAPKAP-K1. This enzyme is activated by phosphorylation in the activation loop. A number of cellular functions of RSK have been proposed, including phosphorylation of transcription factors like CREB and NFκB as well as stimulation of protein biosynthesis by phosphorylation of the ribsomal protein S6.

10.2.2 The JNK/SAPK, p38 and ERK5 MAPK Pathways

Three major MAPK pathways that are activated in response to environmental stress and inflammatory cytokines have been identified in mammals. These pathways have been named according to the terminal MAPK proteins of the core cascade:

- The JNK/SAPK pathway contains the *c-Jun NH_2 terminal kinases* (*JNK*), named alternatively, because of the activation by stress signals, the *stress-activated protein kinases* (*SAPKs*). This subgroup of MAP kinases has, among others, the *transcription factor c-Jun* as substrate.
- A second stress-activated MAPK pathway, the p38 pathway, has the p38 protein as a characteristic terminal kinase (review: Shi and Gaestel, 2002).
- A third, less well-characterized class of stress-activated pathways is the ERK5 pathway, with the ERK5 protein as terminal kinase.

A multitude of input signals can activate these three pathways, and the substrates are very diverse, with substantial overlap in the substrate spectrum, which makes characterization of the pathways difficult. In almost all instances, the stimuli that recruit the JNK/SAPK pathway also recruit the p38 pathway, and these two pathways will therefore be discussed together.

Altogether, the JNK/SAP pathways are characterized by an enormous complexity, and only selected aspects of these pathways can be presented in the following.

Input Signals and Signal Entry Points

External stimuli that activate the JNK/SAPK and p38 pathways include osmotic stress, exposure to bacterial toxins, and environmental perturbations like heat, UV, ionizing

radiation, chemical denaturants like tunicamycin, and alkylating agents. These stresses lead to (among other things) the misfolding of proteins and their accumulation in the endoplasmic reticulum, which in turn induces activation of the JNK/SAPK and p38 pathways.

A variety of MAP3Ks that act upstream of the stress-activated MAPKs have been described, reflecting the many different stimuli that recruit these pathways. The input kinases acting at the level of the MAP3K proteins can be broadly divided into three families: the MEKK1-4 proteins, the mixed lineage kinases (MLKs), and the "thousand and one kinases" TAOs. The latter family is quite specific in activating only the p38 pathway.

The following signaling elements have been shown to recruit the JNK/SAPk and p38 pathways by activating the MAP3Ks:

- **Activation via Rho/Rac proteins**
  Members of this family of small GTPases have members of the family of p21-activated protein kinases (PAKs) as effectors, which phosphorylate and activate MAP3K proteins. Furthermore, the MEKK1 protein has been identifed as a direct effector of GTP-activated Rho/Rac proteins.
- **Activation by tumor necrosis factor receptor (TNFR)**
  The TNFR (see Section 15.5) has associated adaptor proteins that in turn interact with various protein kinases. These include the germinal center kinases (GCKs) that function as MAP4Ks and activate kinases of the MAP3K group, e. g., MEKK1, by a complex mechanism.
- **MAPKs and apoptosis**
  During specific processes of apoptosis, a controlled proteolysis of the MEKK1 polypeptide by caspases (see Chapter 15) is observed. The enzymatically active MEKK1 fragment is thereby released from the multiprotein complexes of the MAPK module, and it is thought that this freely diffusible form of MEKK1 is responsible for its proapoptotic properties.

Substrates of the JNK/SAPK Pathway

As with the ERK proteins, the JNK/SAPK and p38 proteins phosphorylate and activate both transcription factors and other protein kinases (see Fig. 10.4). Some of the protein kinase substrates, the MAPK-activated protein kinases MAPKAP-K2 and -K3, are selectively recruited by stress-activated MAPKs, while others, like the mitogen-activated protein kinase-interacting kinase (MNK), are activated by both stress and mitogenic signals. The MAPKAP-K2 and -K3 polypeptides relay signals to the level of the cytoskeleton; the MNKs regulate the initiation of protein biosynthesis by phosphorylating the initiation factor eIF-4E (see also Section 1.5.5.2).

The transcription factors regulated by the JNK/SAPK and p38 proteins include the Elk1, ATF2, and c-Jun proteins. As expressed by their name, the c-Jun terminal protein kinases phosphorylate the c-Jun protein at residues Ser63 and Ser73. These phosphorylation sites are located within the transactivation domain of c-Jun, and their phosphorylation correlates with enhanced trans-activating activity.

Substrate recognition and selection of the JNK/SAPK and p38 proteins (and also the ERK proteins) are mediated both by specific docking sites and by the nature of the amino acids surrounding the phosphoacceptor site. For the transcription factor substrates, specific docking domains have been identified that are loacted at a distance from the phosphorylation sites in the transactivation domain. These docking sites serve to increase the selectivity and specificity of phosphorylation, and they are used for recruitment of MAPK kinases into protein complexes at promotors, where they can phosphorylate other regulatory transcriptional proteins.

Overall, the picture of the MAPK pathways is complex and incomplete at many points. Parallel pathways exist which may be activated by very different extracellular signals and use different cascade modules. It must be assumed that the different MAPK pathways are crosslinked and that branching may take place at all levels. In addition, the same substrates may be phosphorylated and activated by various MAPK pathways.

Reference

1 Barsyte-Lovejoy, D., Galanis, A. and Sharrocks, A.D. (2002) Specificity determinants in MAPK signaling to transcription factors. *J.Biol.Chem.*, **277**, 9896–9903.

2 Chong, H., Vikis, H.G. and Guan, K.L. (2003) Mechanisms of regulating the Raf kinase family. *Cell Signal.*, **15**, 463–469.

3 Cobb, M.H. and Goldsmith, E.J. (2000) Dimerization in MAP-kinase signaling. *Trends Biochem.Sci.*, **25**, 7–9.

4 Hagemann, C. and Blank, J.L. (2001) The ups and downs of MEK kinase interactions. *Cell Signal.*, **13**, 863–875.

5 Kyriakis, J.M. and Avruch, J. (2001) Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. *Physiol Rev.*, **81**, 807–869.

6 Liebmann, C. (2001) Regulation of MAP kinase activity by peptide receptor signalling pathway: paradigms of multiplicity. *Cell Signal.*, **13**, 777–785.

7 McDonald, P.H., Chow, C.W., Miller, W.E., Laporte, S.A., Field, M.E., Lin, F.T., Davis, R.J. and Lefkowitz, R.J. (2000) Beta-arrestin 2: a receptor-regulated MAPK scaffold for the activation of JNK3. *Science*, **290**, 1574–1577.

8 Raabe, T. and Rapp, U.R. (2002) KSR–a regulator and scaffold protein of the MAPK pathway. *Sci.STKE.*, **2002**, E28.

9 Robinson, M.J. and Cobb, M.H. (1997) Mitogen-activated protein kinase pathways. *Curr.Opin.Cell Biol.*, **9**, 180–186.

10 Shi, Y. and Gaestel, M. (2002) In the cellular garden of forking paths: how p38 MAPKs signal for downstream assistance. *Biol.Chem.*, **383**, 1519–1536.

11 Yordy, J.S. and Muise-Helmericks, R.C. (2000) Signal transduction and the Ets family of transcription factors. *Oncogene*, **19**, 6503–6513.

11
Membrane Receptors with Associated Tyrosine Kinase Activity

Coupling of extracellular signals to tyrosine phosphorylation in the intracellular region may occur by two mechanisms and involves two different receptor types:

- **Receptor tyrosine kinases**
  Ligand binding on the extracellular side is linked to stimulation of tyrosine kinase activity in the cytoplasmic receptor domain for receptors with intrinsic tyrosine kinase activity, the receptor tyrosine kinases (see Section 8.1). The ligand-binding site and the tyrosine kinase are part of one and the same protein.
- **Receptors with associated tyrosine kinases**
  The receptor tyrosine kinases contrast with a group of transmembrane receptors that have no tyrosine kinase activity in the cytoplasmic domain. On ligand binding, this receptor type activates an *associated tyrosine kinase*, so that a signal is created in the form of an intracellular tyrosine phosphorylation. The tyrosine kinase and the receptor are not localized on the same protein in this case. The tyrosine kinase is in most cases permanently associated with the receptor and is activated as a consequence of ligand binding; alternatively it may be located in the cytosol and only bind to the receptor and become activated following ligand binding. Stimulation of the associated tyrosine kinase is then the starting point for transduction of the signal into the interior of the cell. In many cases, mechanisms described in previous chapters are used for the further signal transmission.

11.1
Cytokines and Cytokine Receptors

Cytokines are proteins that serve as signal molecules in cell-cell communication and, as such, perform a central and very diverse function in the growth and differentiation of an organism. Representatives of cytokines control proliferation, differentiation, and function of cells of the immune system and of cells of the blood-forming system. Furthermore, they are involved in the processes of inflammation and in the neuronal, hemapoietic, and embryonal development of the organism. Known cytokines include the interleukins (IL), erythropoietin, growth hormone, interferons (INF), and

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

tumor necrosis factor (TNF) (see Table 8.1). Reviews of cytokines and cytokine receptors are to be found in Ihle et al., 1998 and Moutoussamy et al. (1998).

The cytokines are of considerable medical importance because of their essential function in controlling the immune system, in defense reactions, and for processes of inflammation. Great efforts are therefore being made to elucidate the structure and mechanism of activity of the cytokines and their corresponding receptors and to characterize the components of the signals triggered by cytokines. Many of the cytokines have the character of autocrine hormones, i. e., they only act locally and their targets are cells of the same or similar type as the cytokine-producing cell.

A characteristic that significantly differentiates some of the cytokines from other hormones is the coupling of their activity to cell-cell interactions. The function of some cytokines such as the interleukins IL-4, IL-5, IL-6 and IL-10 is, e. g., closely associated with the interaction between B and T lymphocytes.

11.2 Structure and Activation of Cytokine Receptors

The cytokine receptors are single trans-membrane domain poylpeptides with characteristic extracellular ligand-binding motifs and no known enzymatic activity in their cytoplasmic domains. At least four subtypes of cytokine receptors can be differentiated on the basis of sequence homology (Fig. 11.1). Many members of the cytokine receptors of type 1 regulate growth and transmit mitogenic signals to the cell nucleus. The cytokine receptors of type 2 include the receptors for the interferons α and β. Type 3 includes the receptors for tumor necrosis factor TNF and for CD40 and Fas protein, which are found on T lymphocytes. The latter receptor type will be discussed in Section 15.6.

In the extracellular region, cytokine receptors have characteristic sequence sections that specify the particular receptor type. Cys-rich domains, fibronectin-type III-like domains and immunoglobulin-like domains can be differentiated, among others. Cytokine receptors of type I often have a conserved *WSXWS motif* in the extracellular region. With the exception of ciliary neurotrophic receptor, which is membrane-anchored via a GPI anchor (see Section 3.7.4), all receptors have a single transmembrane domain of 22 – 28 amino acids. Two motifs, called box 1 and box 2, are relatively conserved in the membrane-proximal part of the cytoplasmic domain. However, no known catalytic domain is found on the cytoplasmic part of the receptors.

For some cytokine receptors, including the growth hormone receptor and several interleukin receptors, soluble isoforms have been described that comprise all or part of the extracellular domain and may be able to bind the extracellular ligands. By association with other subunits of heterooligomeric receptors, e. g., the gp130 subunit, these soluble isoforms can function as agonists or antagonists.

The subunit structure of the cytokine receptors is very variable. Three representative cytokine receptor classes will be presented in the following, illustrating the interrelation between receptor subunit structure, ligand binding, and tyrosine kinase activation.

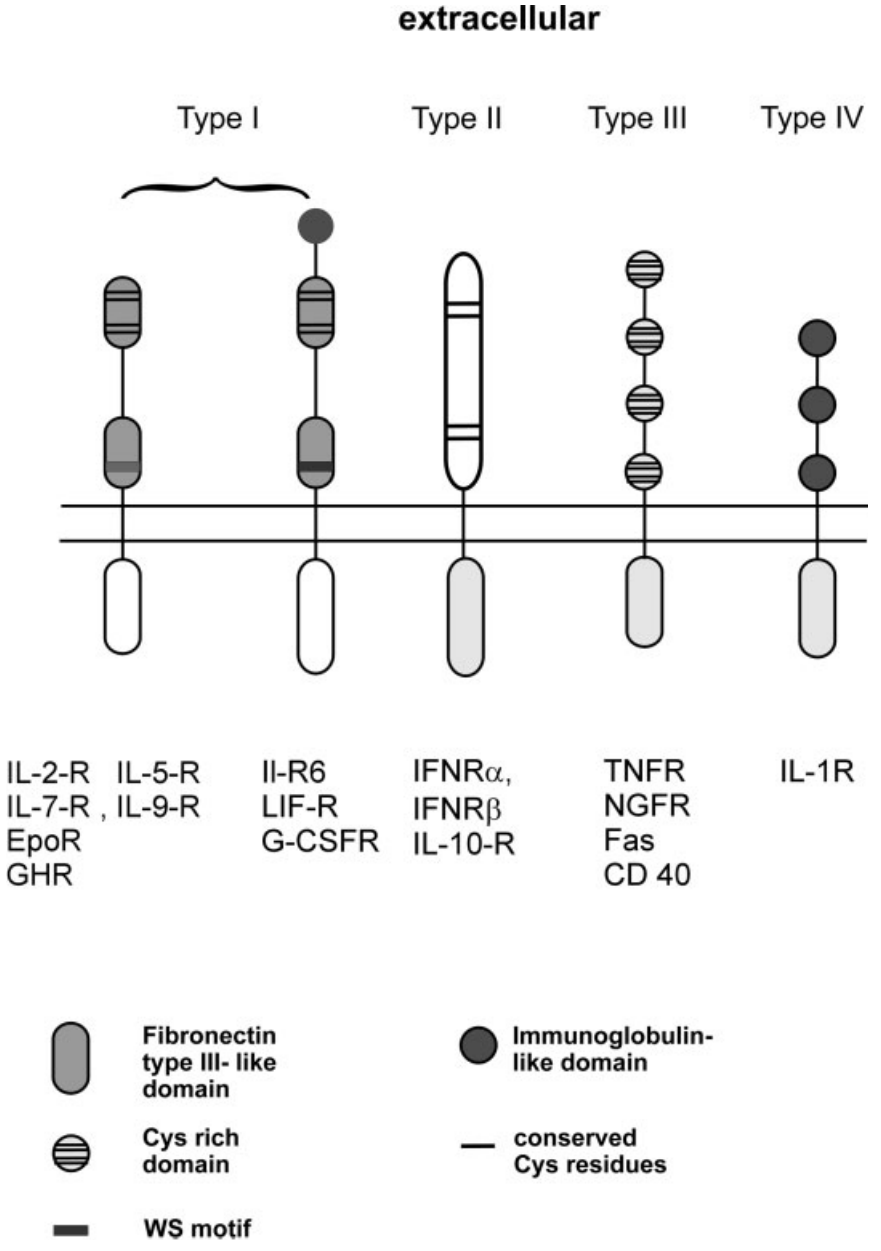


Fig. 11.1 Domain structure of cytokine receptors. Schematic representation of the domain structure of selected cytokine receptors. WS motif: conserved WSXWS sequence (W: tryptophan; S: serine; X: non-conserved amino acid); IL: interleukin; EpoR: receptor for erythropoietin; GHR: growth hormone receptor; LIF-R: leukemia inhibitory factor receptor; G-CSFR: granulocyte colony stimulating factor receptor; IFNR: interferon receptor; TNFR: tumor necrosis factor receptor; NGFR: nerve growth factor receptor; Fas, CD40: transmembrane receptors of lymphocytes.

Erythropoietin Receptor

Only a minority of the cytokine receptors require only one polypeptide for signaling. Examples are the growth hormone receptor and the erythropoietin receptor (review: Constantinescu et al., 2001. These receptors appear to have a dimeric structure in solution, and ligand binding activates the dimeric receptors by a conformation change in the cytoplasmic part of the dimer (Fig. 11.2). As supported by crystallographic studies on liganded and unliganded extracellular domains of the erythropoietin receptor, preformed dimers of this receptor exist in the absence of a ligand, in which the cytoplasmic domains are too far apart to allow trans-activation of the associated tyrosine kinase (review: Frank, 2002). Upon ligand binding, the transmembrane domain and the intracellular domains undergo a conformational change, bringing the associated tyrosine kinase close enough for trans-activation. These studies reveal a great conformational flexibility of the receptor dimer, allowing the same structural parts involved in dimerization also to engage in ligand binding. As illustrated already for the growth hormone receptor (see Fig. 8.5), a monomeric ligand is bound by a dimeric receptor.

Interleukin-6 Receptor, Il-6R

Most cytokine receptors require, for both high-affinity ligand binding and signaling to the cytoplasm, the interaction with one or more additional subunits. The interleukin-6 receptor (review: Kerr et al., 2001) exemplifies receptor types which use different subunits for ligand binding and signal transmission to the cytoplasm. Binding and recognition of the ligand Il-6 is accomplished by the α-subunit of the functional receptor, which has only a short cytoplasmic tail and is therefore not able to conduct by its own the signal into the cytoplasm. Rather association of another subunit, the gp130 subunit (review: Arzt, 2001), is required for signal transmission to the cytoplasm (Fig.11.3). Complex formation between Il-6 and Il-6α triggers dimerization of the gp130 subunit, which has a tyrosine kinase permanently associated. In the oligomeric complex formed, the tyrosine kinase is activated and the signal is transduced further. *Two separate polypeptide chains* are needed for ligand binding and communication with the cytosolic side in this case. One serves for specific ligand binding, and the other serves to pass the signal into the cell interior. Even a soluble form of Il-6α, comprising only the extracellular part, can bind the ligand and activate surface-expressed gp130 for Il-6 signaling. The gp130 subunit is also used for signaling by other ligands, e. g., the leukemia inhibitory factor (LIF) and others. These ligands bind to their specific α-subunit and use the gp130 subunit as a common signal transducer.

Il-2 Receptor, Il-2R

Signaling by Il-2 (review: Ellery and Nicholls, 2002) requires three receptor subunits, Il-2Rα, Il-2Rβ, and Il-2Rγc (Fig. 11.4). The α-subunit has primarily the function of an affinity modulator. Its presence is required for high-affinity binding of Il-6 by the other two subunits, which bind Il-6 in the absence of Il-2Rα only with intermediate affinity. Il-2Rβ and Il-2Rγc together are necessary and sufficient for efficient signaling, and both connect ligand binding to the activation of intracellular signaling components. Despite an absolute requirement of Il-2Rγc for signaling, the majority of downstream-signaling pathways link through the Il-2Rβ subunit to the activated receptor.

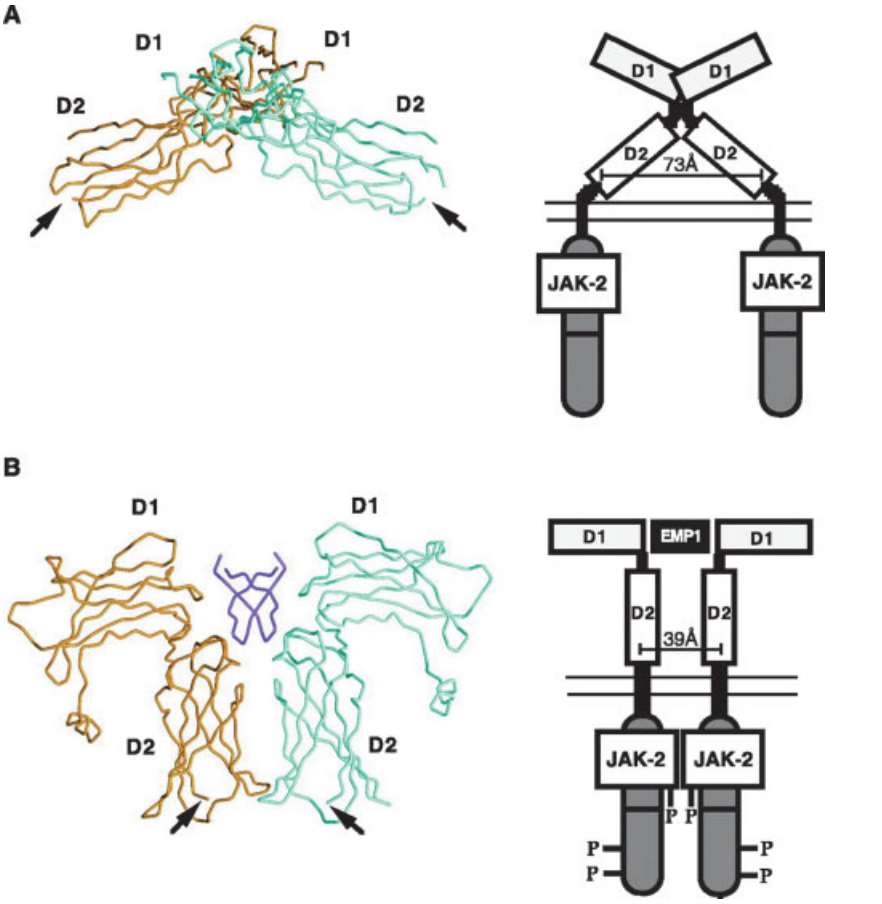


Fig. 11.2 Comparison of the liganded and unliganded erythropoietin receptor (EPOR) dimer configurations (from Livnah et al., 1999). The structure of a dimer of a EPOR fragment comprising part of the extracellular ligand binding domain (residues 1 to 225 with subdomains D1 and D2) of the human EPOR is shown in the absence (A) and presence (B) of a peptide (EMP1) that mimics the erythropoietin ligand and functions as an agonist. The two EPOR fragments are shown in cyan and gold, with their subdomains labeled D1 and D2. The two EMP1 peptides are shown in purple and the membrane-proximal ends of the D2 domains are indicated by a black arrow.

A) Without ligand, two EPOR fragments form a cross-like self dimer. The D1 domains of each monomer point in opposite directionsand the membrane-proximal ends of the D2 domains are separated by 73 Å. In the schematic of the unliganded self dimer (right) , the scissors-like dimer configuration keeps the intracellular ends far enough apart such that autophosphorylation of the intracellularly associated tyrosine kinase Jak-2 cannot occur and hence other phosphorylation events, such as on the cytoplasmic domain of EPOR do not occur.

B) Two EMP1 peptides bind to the EPOR fragments in a symmetrical manner. EMP1 (or eythropoietin) induces a close dimer association of both D1 and D2 subdomains so that the ends of the D2 domains are much closer as compared to the unliganded situatuion. As a consequence, the intracellular regions become substrates for phosphorylation by two Jak-2 molecules (right).

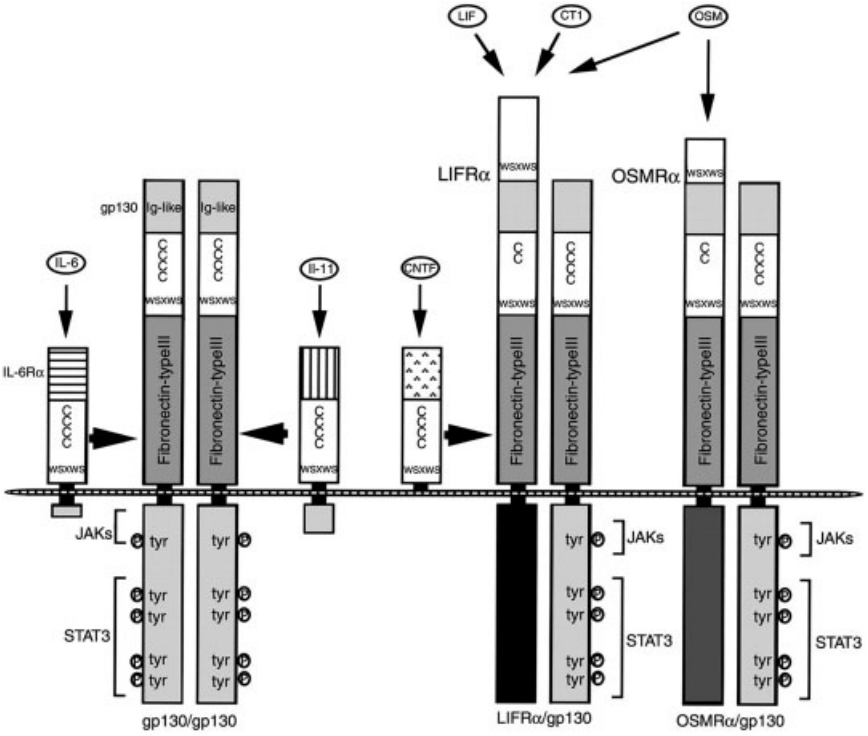


Fig. 11.3 Schematic model of the IL-6/gp130 receptor system. The specific cytokine-binding subunits and gp130 belong to a cytokine receptor family characterized by four positionally conserved cysteine and a WSXWS motif. Functional receptor complexes are induced by the different cytokines: IL-6 and IL-11 induce gp130 homodimerization.; CNTF, LIF, CT-1 and OSM induce LIFRα/gp130 heterodimers; and OSM may also induce OSMRα/gp130 heterodimers. Following ligand binding, gp130 is responsible for signal transduction through the JAK/STAT pathways.
CT-1: cardiotrophin 1; CNTF: ciliary neurotrophic factor; LIF: leukemia inhibitory factor; OSM: oncostatin M.

Nearly all signals that are generated upon Il-2R activation can be mapped to the cytoplasmic tail of the Il-2Rβ subunit. The γc subunit is shared with other cytokine receptors, e. g., the receptors for IL-4, IL-7 and IL-9 and is therefore – like gp130 – a common cytokine receptor subunit.

Conformational changes in preexisting oligomers and/or a ligand-induced heterooligomerization of receptor subunits have emerged as a general principle underlying cytokine receptor activation, a situation that is similar to the activation of the receptor tyrosine kinases (Chapter 8). No tyrosine kinase activity or other enzyme activity has been found in the intracellular sequence sections of the cytokine receptors. Rather, ligand binding is linked to activation of an associated tyrosine kinase, which then conducts the signal further without the activated receptor actually performing any enzyme function (see Section 11.2.1).

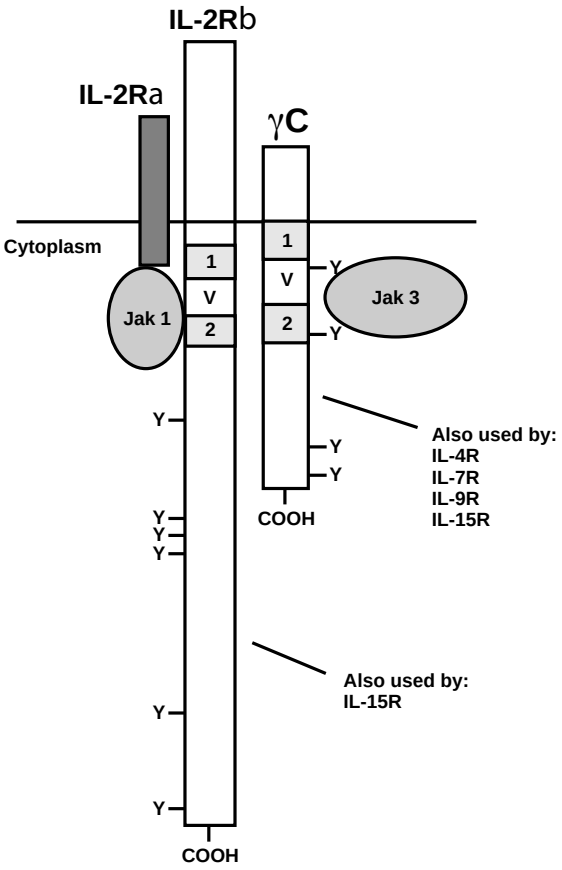


Fig. 11.4 Schematic diagram of the Il-2 receptor complex. The Il-2R is composed of three subunits, IL-2Rα, IL-2Rβ γc. The Box 1 and Box 2 domains are conserved among members of the cytokine receptor superfamily, and a variable spacer region V intervenes Box1 and Box2. The Box1-V-Box2 regions encompass the sites of association with the Janus kinases (Jak). Jak1 associates constitutively with IL-2Rβ and Jak3 associates constitutively with γc. Both IL-2Rβ and γc encode multiple cytoplasmic tyrosine residues with approximated locations shown. At least some of these tyrosines on both chains become phosphorylated following receptor activation whereupon they serve to recruit effector proteins to direct downstream signaling events.

11.2.1 Activation of Cytoplasmic Tyrosine Kinases

As a consequence of ligand binding to cytokine receptors, activation of a tyrosine kinase activity, which is not part of the receptor protein, is observed. In most cases this tyrosine kinase is permanently associated with one of the subunits of the receptor, and ligand-induced restructuring or hetero-oligomerization of the receptor induces activation of this tyrosine kinase. The tyrosine kinases most frequently associated with cytokine receptor subunits belong to the family of Janus kinases. However, protein tyrosine kinases of the Src family like Lck, Fyn and Tyk have also been shown to be direct downstream components of the cytokine receptors (see Section 11.2.2).

The following steps are involved in cytokine receptor signaling (Fig. 11.5):

- **Activation of the associated tyrosine kinase by autophosphorylation**
  This step occurs immediately after ligand binding and is thought to function via autophosphorylation in trans between neighboring tyrosine kinases associated with

the oligomeric receptor. This activation involves phosphorylation in the activation loop of the tyrosine kinase.

- **Phosphorylation of receptor subunits**
  The activated tyrosine kinase then phosphorylates tyrosine residues in the cytoplasmic region of the receptor subunits. The phosphotyrosine residues serve as attachment points for the recruitment of other signaling proteins.
- **Binding of other signaling proteins**
  Distinct phosphotyrosine residues are used for the attachment of signaling proteins that carry phosphotyrosine-binding domains like SH2 or PTB domains. In this way,

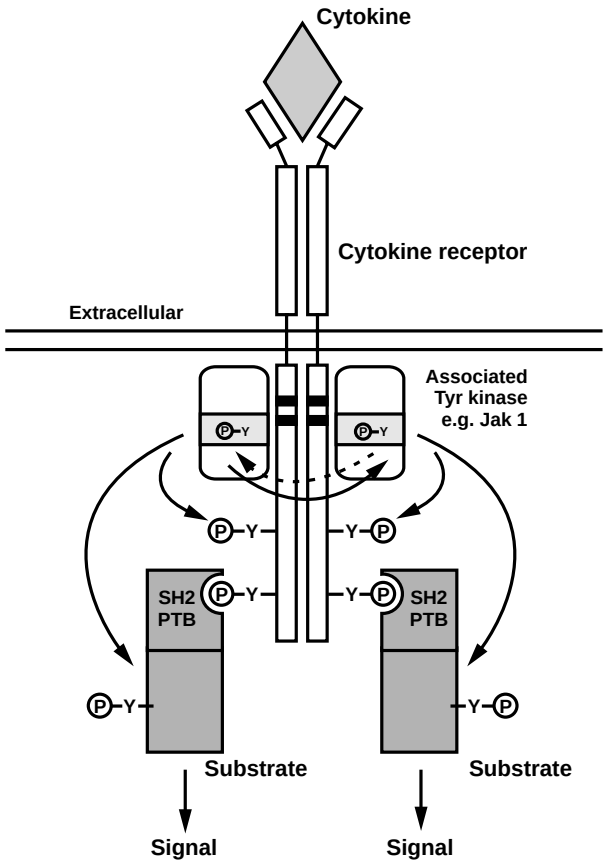


Fig. 11.5 Steps involved in cytokine signaling. Binding of a cytokine to the extracelluar ligand binding domain of the cytokine receptor activates the tyrosine kinase activity of the associated tyrosine kinase (e. g. Jak1). Tyrosine phosphorylation occurs in trans between neighbouring kinase molecules. The activated protein kinase also catalyzes Tyr phosphorylation of the cytoplasmic domain of the receptor. The phosphotyrosine residues serve as attachment points for adaptor proteins or other effector proteins containing phosphotyrosine-binding motifs (PTB or SH2). The signal is then transmitted further into the cytoplasm.

adaptor proteins like Shc or IRS are recruited to the cytokine receptor. As a consequence, signals are passed on to a variety of intracellular signaling pathways.

- **Phosphorylation of substrates**
  Signaling proteins that have been recruited to the activated receptor are often substrates for phosphorylation by the activated tyrosine kinase. Examples are the Stat proteins (see below), the Shc adaptor protein, and PI3 kinase. Il-2Rβ activates PI3 kinase by inducing phosphorylation on tyrosine residues of the p80 regulatory subunit and recruiting PI3 kinase to the cell membrane.

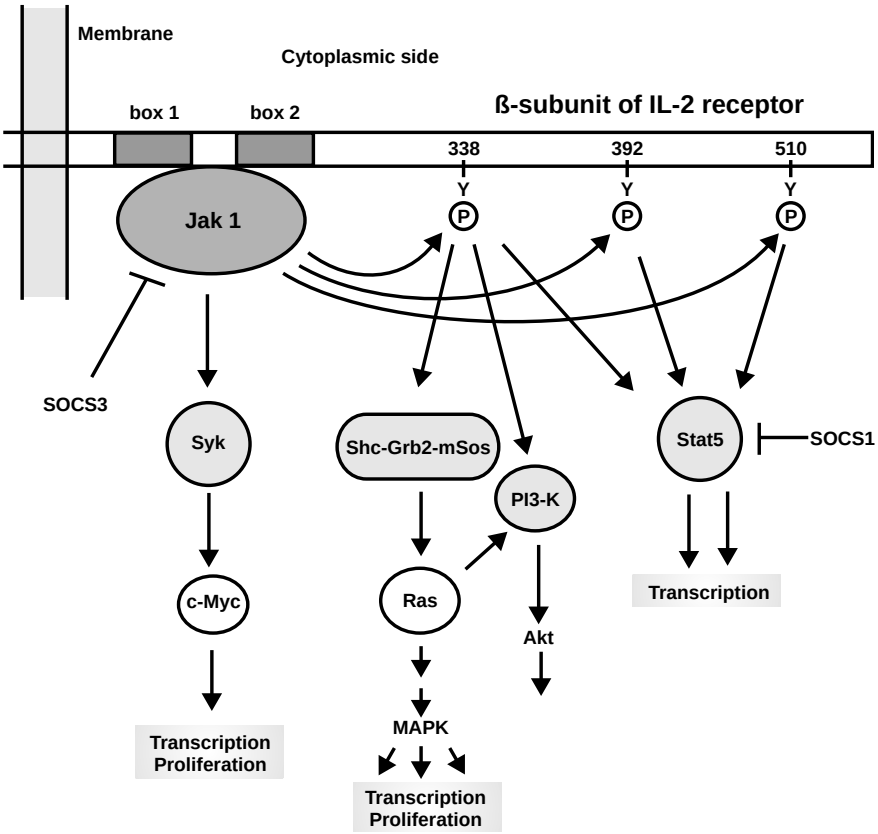


Fig. 11.6 Signaling by the β-subunit of the IL-2 receptor. The cytoplasmic portion of the β-subunit is shown with tyrosine phosphorylation sites that mediate further downstream signaling events. Jak1 is constitutively associated with the Box1 and Box2 regions. Following ligand binding, Jak1 becomes activated and phosphorylates distinct tyrosine residues on the β-subunit which serve as attachment points for SH2 domains of further signaling proteins as indicated. The Stat5 transcription factor can bind to all three phosphorylation sites, is phosphorylated by Jak1and then translocates to the nucleus to activate transcription of target genes. Jak1 also phosphorylates and activates the nonreceptor tyrosine kinase Syk which has the transcription factor Myc as a substrate. Inhibition and attenuation of IL-2Rβ signaling is exerted by the SOCS proteins which can inhibit Jak1 or interfere with Stat activation, among others. PI3-k: PI3-kinase; SOCS: suppressor of cytokine signaling; Akt: Akt kinase.

The principles of cytokine receptor signaling are illustrated in Fig. 11.6 for the β-subunit of the IL-2 receptor. Il-2R is composed of the Il-2Rα, Il-2Rβ, and Il-2Rγc subunits. Of these, only the Il-2Rβ and Il-2Rγc subunits conduct the signal further into the cytoplasm. Two protein tyrosine kinases, Lck kinase and Jak1, have been found to be associated with the Il-2Rβ polypeptide. Most specific signals are mediated by Jak1, while the *in vivo* function of the Lck protein in Il-2R signaling is not yet fully established. Another member of the Janus kinase family, Jak3, is firmly associated with the Il-2Rγc subunit. It is not yet clear how the different kinases cooperate in Il-2R signaling. Most signaling specificity is imparted by Jak 1, which phosphorylates specific tyrosine residues on Il-2Rβ. Phosphorylation of Tyr338 by Jak1 induces binding of the adaptor Shc and subsequently the recruitment of the Grb-mSos complex. As a result, the Ras protein is activated and a signal is delivered to the effector pathways of the Ras protein, and activation of the MAPK/ERK pathway has been observed many times as a consequence of Il-2R activation. Jak1-mediated phosphorylation of tyrosine residues (e. g. Tyr392 and Tyr510) in the C-terminal half of the cytoplasmic tail creates binding sites for the Stat5 protein, which is then phosphorylated by Jak1 and activated for further signaling. Another effector is the PI3-kinase, which binds to the above-mentioned phosphotyrosine residues through the SH2 domain of its p85 regulatory subunit.

Starting from the activated tyrosine kinase, the signal may be conducted along different signaling pathways, depending on the receptor type:

- Jak-Stat pathway (see below)
- Ras pathway
- MAP kinase pathway
- protein kinase C via phospholipase Cγ
- PI3-kinase pathway.

Negative Regulation of Cytokine Receptor Signaling: The SOCS Proteins

The activity of the cytokine receptors is under tight control to avoid prolonged activation and excessive signaling. Several mechanims exist by which termination and damping of cytokine receptor signaling is achieved:

- recruitment of tyrosine phosphatases (see Section 8.4)
- endocytosis and targeted degradation via the ubiquitin-proteasome pathway (see Section 2.6)
- recruitment of inhibitory proteins: SOCS proteins (review: Krebs and Hilton, 2001).

A family of proteins has been identified whose members act as negative regulators of cytokine signaling These proteins have been named suppressors of cytokine signaling (SOCS proteins), and up to now at least five subfamilies have been identified within the SOCS family. The SOCS proteins share a conserved sequence motif, the SOCS box, plus either an SH2 domain or other domains capable of mediating protein-protein interactions. The mRNAs coding for the SOCS proteins are induced in response to stimulation with a large number of cytokines, and negative regulation of cytokine signaling by the SOCS family members has been clearly shown. Thereby, SOCS pro-

teins can act in a classical negative feedback loop: they inhibit the signaling pathway that stimulates their own production. The transcription of the *socs1* gene is, e. g., enhanced by the STAT transcription factors, which are major players in cytokine signaling and are commonly activated during cytokine signaling (see below).

The mechanisms by which SOCS family members inhibit cytokine receptor signaling are diverse and depend on the nature of the SOCS protein and the receptor involved. The SOCS1 protein, e. g., has been shown to associate with a phosphotyrosine residue of the activation loop of the JAK1 tyrosine kinase, leading to inhibition of the kinase activity (see Fig.11.6). Another way by which SOCS proteins are thought to inhibit cytokine signaling is by competition with STAT proteins for phosphotyrosine-binding sites on receptor subunits.

11.2.2 The Jak-Stat Pathway

The Jak-Stat pathway is a signaling pathway, starting from cytokine receptors, that allows a very direct signal transduction from the membrane to the cell nucleus using only a few coupling elements. Many cytokines use this pathway to bring about a rapid change in the transcription activity of specific gene sequences (reviews: Moutoussamy et al., 1998; Imada and Leonard, 2000).

11.2.2.1 The Janus Kinases

The family of protein kinases most often involved in signal transduction via cytokines includes the *Janus kinases* (*Jak kinases*). At least four different Jak kinases are known in mammals (Jak1, Jak2, Jak3 and Tyk2). A characteristic feature of the structure of Jak kinases is the occurrence of two tyrosine kinase domains (Fig. 11.7): the JH1 and JH2 domains. However, only JH1 possesses all of the structural features considered necessary for a functioning kinase activity. JH2 is also called a pseudokinase domain, and its function is still unknown.

Most Jak kinases are constitutively associated with a cytoplasmic section of the receptor, which is in the vicinity of the membrane and contains two conserved sequence elements, Box 1 and Box 2. Cytokine-induced conformational changes of the receptor oligomer bring about a change in the juxtaposition of the associated JAKs, allowing cross-phosphorylation of neighboring kinases, which leads to their activation (see Fig. 11.5). Activation of the Jak kinases may take place in a homodimeric receptor complex or it may also occur in hetero-oligomeric complexes.

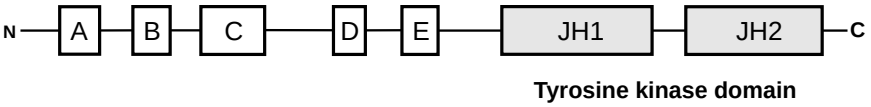


Fig. 11.7 Domain structure of the Jak kinases. JH1 is the catalytic tyrosine kinase domain. JH2 shows similarity to a tyrosine kinase domain. The domains A - E are homologous elements of the Jak kinase family. JH: Janus kinase homology region.

11.2.2.2 The Stat Proteins

Starting from the activated Jak kinases, a signaling pathway leads directly to transcription factors that are phosphorylated by the Jak kinases on tyrosine residues and activated for stimulation of transcription (reviews: Horvath and Darnell, 1997; Bromberg, 2001). These transcription factors belong to a class of proteins known as Stat proteins (Stat = signal transducer and activator of transcription). At least seven different Stat proteins are known (Stat1, Stat2, Stat3, Stat4, Stat5a, Stat5b, and Stat6). The first Stat proteins, Stat1 and Stat2, were found in association with signal transduction via interferon γ. Figure 11.8A shows the domain organization of Stat proteins, and the structure of Stat1 bound to DNA is depicted in Fig. 11.8B.

The Stat proteins are found in a latent form in the cytosol and are activated by phosphorylation on a tyrosine residue located ca. 700 residues from the N-terminus. This phosphorylation can be catalyzed by Jak kinases or by other tyrosine kinases (see below). The signaling pathway where the STAT proteins are activated via Jak kinases is named the JAK/STAT pathway. On binding of the cytokine to the receptor and activa-

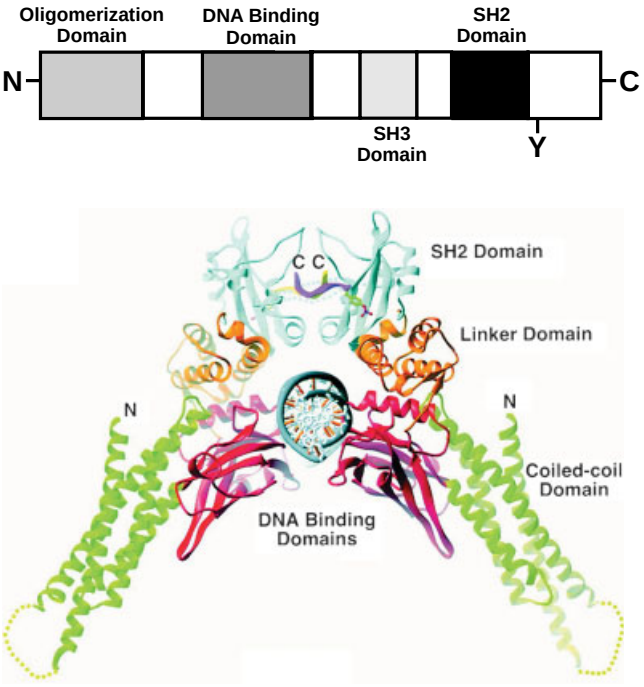


Fig. 11.8 A) Domain structure of STATs. STATs bind to receptors and dimerize via bivalent SH2-phosphotyrosine interactions. Phosphorylation of the conserved tyrosine is required for STATs dimerization. The N-terminal region mediates oligomerization of STAT dimers. B) Dimer of Stat2 bound to DNA. Dimerization is mediated by reciprocal SH2-Tyr-phosphate interactions of the monomers. The view is along the DNA helix. The DNA binding domain is in red, the linker domain in orange and the SH2 domain in light green. The C-termini of the two Stat2 molecules are shown in yellow and magenta.

tion of the Jak kinase, the Stat proteins are recruited, via their SH2 domains, to phosphotyrosine residues of the receptor-kinase complex and are then phosphorylated by the Jak kinase on a conserved Tyr residue (Tyr701 for STAT1) at the C-terminus.

In the unphosphorylated form, the Stat proteins exist as monomers, whereas in the phosphorylated form, they form homodimers, heterodimers or higher oligomers and are transported as such into the nucleus (Fig. 11.9), where they act as transcriptional activators.

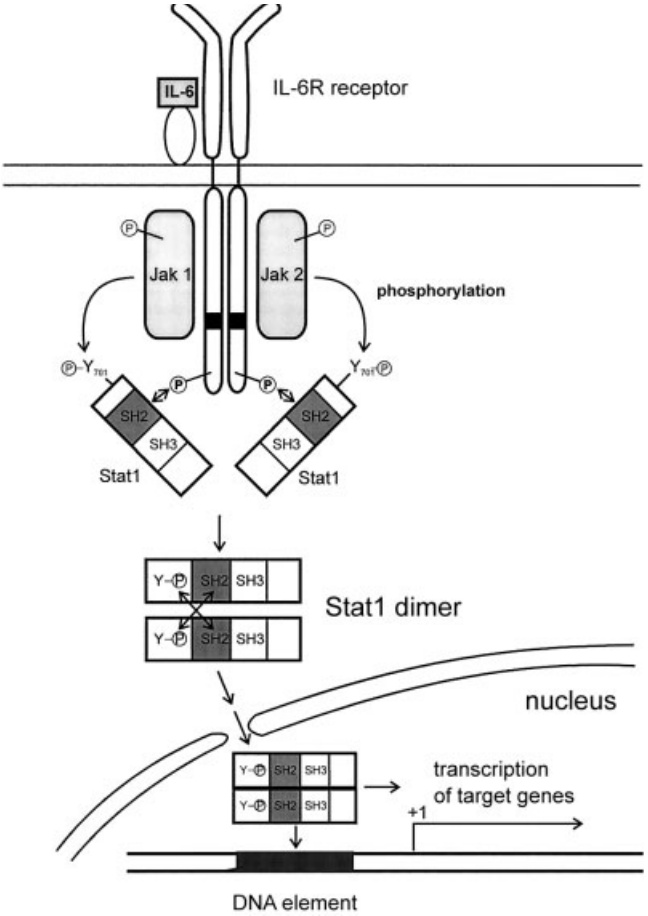


Fig. 11.9 Model of activation of Stat proteins. The Stat proteins are phosphorylated (at Tyr701 for Stat1) as a consequence of binding to the receptor-Jak complex, and Stat dimers are formed. The dimerization is mediated by phosphotyrosine-SH2 interactions. In the dimeric form, the Stat proteins are transported into the nucleus, bind to corresponding DNA elements, and activate the transcription of neighboring gene sections. In the figure, activation of Stat proteins is shown using the IL-6 receptor as an example. Other Jak kinases and Stat proteins may also take part in signal conduction via IL-6, in addition to the Jak kinases and Stat1 shown.

The STAT proteins have an N-terminal oligomerization domain, an SH3 and an SH3 domain, a DNA binding domain and a C-terminal transactivation domain. Dimerization is mediated by the phosphotyrosine residue and the SH2 domain. Highly resolved structural investigation show that the phosphotyrosine residue of one Stat protein binds to the SH2 domain of the partner and vice versa, so that the phosphotyrosine-SH2 bonds function as a double clasp (structure in complex with DNA: Becker et al., 1998). The binding to DNA is in the form of a dimer, with the Stat-DNA

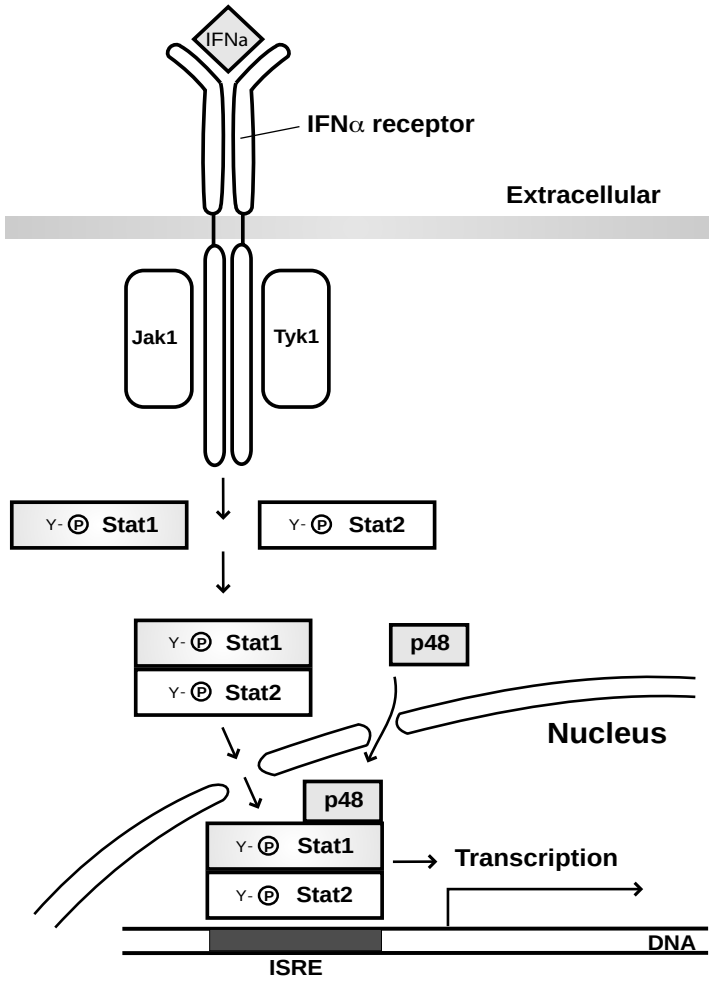


Fig. 11.10 Scheme of signal transduction via interferon α. The receptor for interferon α (IFNα) binds and activates the Jak kinases Jak1 and Tyk1. These phosphorylate the Stat factors Stat1 and Stat2, leading to formation of Stat1-Stat2 heterodimers. The heterodimers are transported into the nucleus and bind to a corresponding DNA element known as ISRE (interferon stimulated response element). Another protein, p48, is also involved in transcription activation of the interferon regulated gene.

complex showing a large similarity to the structure of the NFκB-DNA complex (see Fig. 11.8B). In the nucleus, they bind to corresponding DNA elements in the promotor region of cytokine-responsive genes. Stat-binding sites are often arranged in tandem on promotors, and STAT tetramers are then formed on the DNA. In the course of transcription activation, STAT proteins make contacts with the RNA polymerase II machinery via the transactivation domain. Furthermore, STAT proteins interact with and recruit histone actelyase complexes (see Section 1.14.6), and they often cooperate with other transcription factors such as glucocorticoid receptors and c-jun within enhanceosome complexes.

The Jak-Stat pathway for the interferons α, β and γ has been well investigated. The interferon γ pathway uses Stat1 protein, which binds as a homodimer to the corresponding GAS element (GAS: interferon γ activation site). The GAS element includes the consensus sequence TTNCNNNAA and is found in a modified form in many cytokine-controlled gene sections. The signaling pathway of the interferons α and β uses the Stat1 and Stat2 proteins in particular; binding to the corresponding DNA-binding element also requires the help of a 48 kDa protein (Fig. 11.10).

The Jak-Stat signal transduction is an example of a signaling pathway in which a signal is coupled, in the form of a tyrosine phosphorylation, directly to activation of a transcription factor. In contrast to other signaling pathways that also regulate transcription processes, e. g., the Ras/MAPK pathway, the Jak-Stat pathway is impressive in its simple concept and the small number of components involved.

The Jak/STAT pathway is not the only pathway by which STAT proteins transmit signals to the level of transcription. Over 40 different proteins have been identified that induce phosphorylation and activation of STAT proteins, and several signaling pathways converge at the level of the STAT proteins. Stat proteins can be activated via
- Janus kinases, as described
- receptor tyrosine kinases such as EGF-R and PDGF-R
- nonreceptor tyrosine kinases (e. g. Src kinase and Abl kinase, Section 8.3)
- G-protein-coupled receptors via indirect ways.

11.3 T and B Cell Antigen Receptors

At the surface of T and B lymphocytes, specific receptors are found that bind antigens and set intracellular signal chains in motion (review: Qian and Weiss, 1997). These may lead to increased cell division, programmed cell death, or a functional recoining of lymphocytes.

The receptors of the B lymphocytes recognize antigens in the form of foreign proteins, which exist in soluble, particle-bound or cell-bound forms.

The receptors of the T lymphocytes, in contrast, recognize antigens only in the course of a cell-cell interaction between the T lymphocyte and an antigen-presenting cell. The antigen-presenting cell presents the processed (i. e., proteolytically digested to form peptides) foreign protein as a peptide. The peptide is bound to the MHC complex (MHC: major histocompatibility complex) of the antigen-presenting cell and is recog-

nized in this form by the receptor of the T lymphocyte. The MHC proteins are transmembrane proteins that exist as heterodimers and possess a binding site in the extracellular region for the antigen to be presented (= processed peptide). B and T cell antigen receptors bind the antigen under very different conditions. Despite this, the binding event triggers very similar signal chains and initiates similar reactions.

11.3.1 Receptor Structure

The T and B cell antigen receptors are generally composed of several subunits, whereby the functions of ligand binding and conduction of the signal are localized on separate subunits (reviews: Garcia et al., 1999, Germain, 2001).

The T cell antigen receptor contains a minimum of 8 polypeptides and may be divided into two functional complexes, in which one mediates ligand binding and the other performs the signal transduction. Antigen binding takes place via the Ti-α and β subunits, which only have very short cytoplasmic structural portions and are not di-

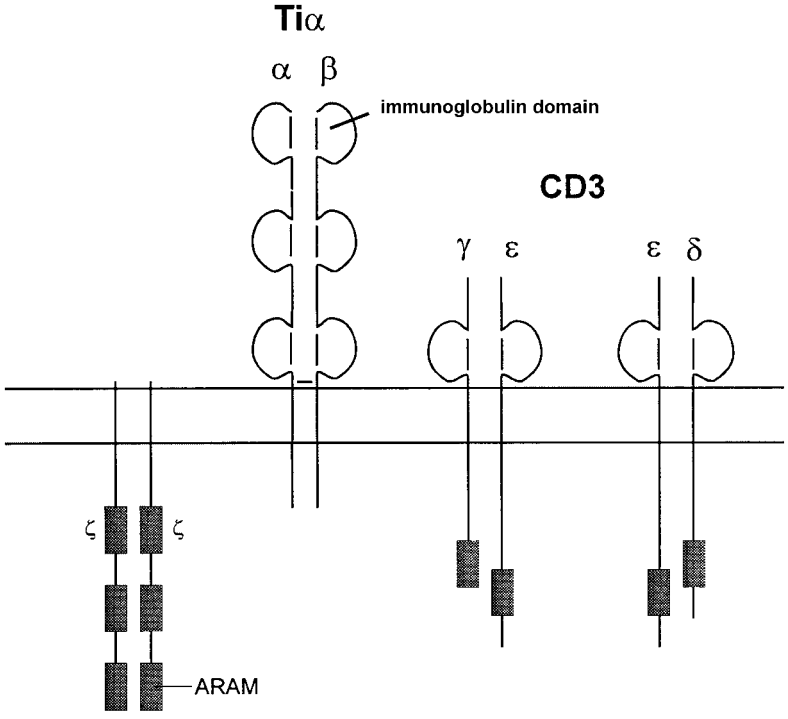


Fig. 11.11 Subunit structure of the T cell receptors. The figure shows the different subunits of T cell receptors in a highly simplified representation. The stoichiometry of the subunits in the complete receptor is not clear. The $\alpha\beta$ chains are also known as the Ti$\alpha\beta$ complex; the $\gamma\varepsilon$ and $\delta\varepsilon$ chains together form the CD3 complex. ARAM: antigen recognition activation motif.

rectly involved in conduction of the signal on the cytosolic side. The function of signal conduction is performed by other polypeptide chains, namely the CD3γ, CD3δ and CD3ε chains and the ζ chain (Fig. 11./11). These have characteristic sequence motives on their cytoplasmic side, which are required for conduction of the signal further to downstream protein tyrosine kinases. The sequence motifs include two pairs of Tyr and Leu residues in the consensus motif (D/E)XXYXXL(X)$_{6\text{-}8}$YXXL and are known as the *antigen recognition activation motif* (ARAM) or also as the *immunoreceptor tyrosine activation motif* (ITAM). ITAM motifs are found in the CD3γ, CD3δ and CD3ε chains and in the ζ chain. ITAM motifs also occur in B cell receptors. Following TCR stimulation, the tyrosine residues within the ITAMs become phosphorylated by Src family tyrosine kinases, a key early event in the TCR-signaling cascade. Cooperation with other receptors that may help in a synergistic manner to trigger a signal is a particular feature of signal transduction via T and B cell antigen receptors. These other receptors are known as *coreceptors*. Examples are the CD4 and CD8 proteins, which are involved in activation of T cell antigen receptors. The coreceptors are essential for signal transduction and are involved in binding and activation of downstream tyrosine kinases such as Src kinase. Furthermore, they have an amplifying effect on the sensitivity and specificity of antigen binding.

11.3.2 Intracellular Signal Molecules of the T and B Cell Antigen Receptors

The T and B cell antigen receptors do not have any intrinsic protein tyrosine kinase activity. Rather, antigen binding leads to activation of protein tyrosine kinases of the Src family (Lck, Fyn kinases) on the cytoplasmic side of the receptor. These kinases are thought to be constitutively associated with the TCR complex and become activated upon ligand binding by a mechanism still to be resolved. Following this early activation step, the ITAM motifs are phosphorylated and serve as docking sites for the binding of kinases of the Syk family (Syk, ZAP70 kinase). ZAP70 kinase (ZAP = ζ associated protein 70) binds with the help of its SH2 domains to the phosphotyrosine residues thus created, and it is then activated by phosphorylation. The next step in the signaling cascade is thought to be the ZAP70-catalyzed phosphorylation of an adaptor protein, named LAT (Linker for activation of T-lymphocytes), which is a transmembrane protein found in glycolipid-enriched subdomains of the cell membrane. From that point on, various other signaling proteins are recruited, either directly or indirectly, into TCR signaling by binding via SH2 domains to phosphotyrosine residues of LAT (see Fig. 11.12). These signaling proteins include PLC-γ and adaptor proteins such as Grb2, which assemble further proteins like GEFs (mSos, Vav) and nonreceptor tyrosine kinases (Tec) into the TCR-signaling cascade.

An inhibitory phosphorylation of the tyrosine kinases Lck and Fyn is also important for signal transduction, and this is performed by Csk kinase. This phosphorylation takes place at the C-terminus of Lck and inhibits the kinase activity in a similar way to that already structurally illustrated for Src kinase (see Section 8.3.2).

As a consequence, the following signaling pathways are activated following stimulation of T cell receptors:

– Ras pathways with MAPK cascades
– hydrolysis of PtIns phosphates: via phospholipase Cγ
– Ca^{2+}-signaling pathways: via phosphorylation of the $InsP_3$ receptor and via PLCγ.

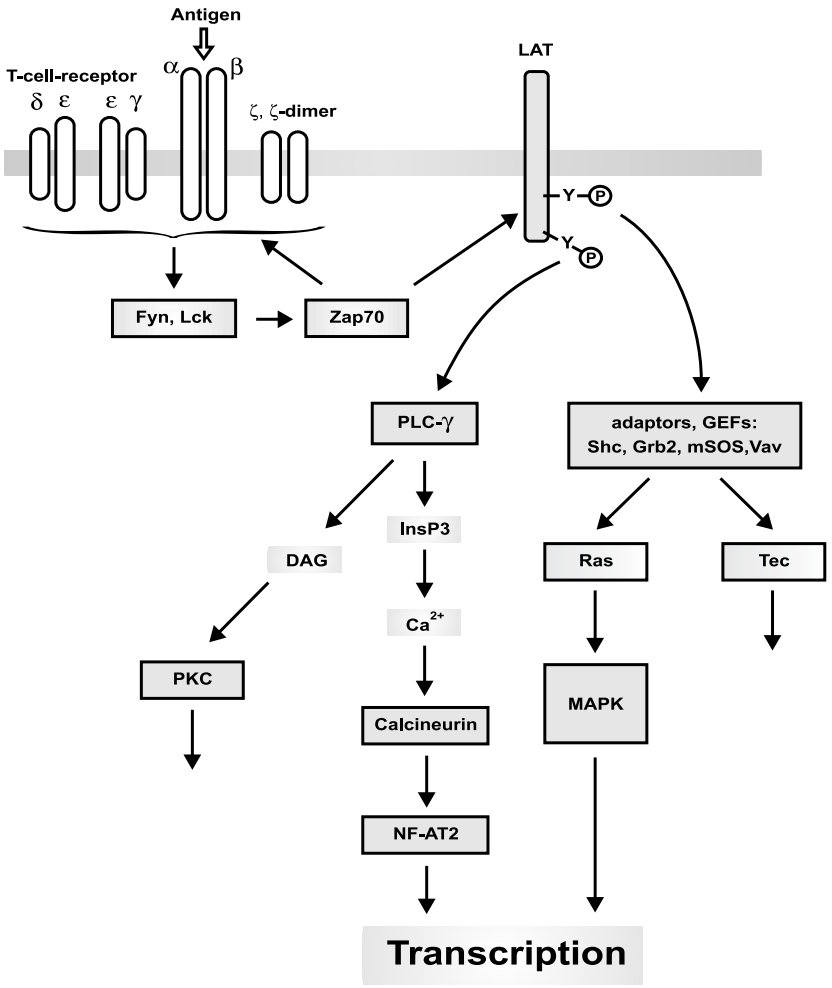


Fig. 11.12 Overview of signaling pathways associated with activation of lymphocytes. The triggering signal for activation of T lymphocytes is generally antigen binding to the T cell receptor. The activated receptor passes the signal on to associated tyrosine kinases like Fyn, Lck and ZAP70. These phosphorylate the transmembrane protein LAT on cytoplasmic tyrosine residues. The LAT phospho-tyrosine residues are docking sites for adaptors (Shc, Grb2) and GEFs which pass a signal to the Ras/MAPK pathway or to the nonreceptor tyrosine kinase Tec. Furthermore, PLCγ is recruited to LAT and a Ca^{2+}/ diacylglycerol signal is generated leading - via calcineurin/NF-AT2 (see also Fig. 7.19) - to transcription of target genes and to activation of protein kinase C (PKC). DAG: diacylglycerol; InsP3: inositol 1,4,5 trisphosphate.

11.4 Signal Transduction via Integrins

The structure and function of the cell formations of higher organisms are highly dependent on adhesive interactions based on direct cell-cell contact and on interactions of cells with the extracellular matrix. Adhesion between cells and with the extracellular matrix has a regulatory influence on migratory behavior, proliferation, and differentiation of an individual cell within the cell formation. The adhesion processes thus do not serve simply to hold cells together in the formation. They also have a regulatory effect on central physiological functions of a cell. Surface receptors that can specifically bind to a neighboring cell or to the extracellular matrix serve as mediators of adhesion processes, and, as a consequence, intracellular signaling pathways are activated. The protein family of the integrins are one such group of surface receptors which are of central importance (reviews: Howe et al., 1998; Schwartz, 2001).

The integrins define attachment points for the extracellular matrix and for contact with neighboring cells, and they are involved in signal transduction into the cell interior. With these functions, the integrins are involved in the regulation of embryonal growth, tumor formation, programmed cell death, tissue homeostasis, and many other processes in the cell. The integrin receptors (usually just called integrins) have a special position amongst the transmembrane receptors. The integrins form a transmembrane bridge between extracellular scaffold structures and the intracellular cell scaffold. They also convert cell-cell contacts, interactions with the extracellular matrix (ECM), and mechanical stress into intracellular signal processes and vice versa.

An important function of integrins is the linkage between the extracellular matrix and the intracellular actin cytoskeleton. This bonding often takes place in structures known as *focal adhesion points*. These are multiprotein complexes in which specific contacts are formed on the cytosolic side between the integrin receptors, proteins of the intracellular matrix, and actin filaments (Fig. 11.13). Focal adhesion points form linkages between morphological structures of the cell and signal transduction pathways. These multiprotein complexes have an attachment point for the actin cytoskeleton and binding sites for signal proteins.

The integrin receptors are made up of α and β chains, which each have a transmembrane element. At least 16 different α chains and 8 different β chains are known at present. The integrin receptors form heterodimers, which is why there is such a large structural and functional diversity of integrins. The extracellular ligands of the integrins are mostly components of the extracellular matrix such as fibronectin and collagen. These are multivalent ligands immobilized on fibrillar structures. Extracellular ligands may, however, also be soluble proteins or surface proteins of neighboring cells.

The ligands of the integrin receptors are generally multivalent, and their binding leads to crosslinking and clustering of integrin receptors. By this process, the integrin receptors are activated for further signal transduction.

On the cytoplasmic side, a multitude of proteins have been reported to be implicated in integrin signal transduction. These proteins either interact directly with the cytoplasmic tails of the α and β subunits or they are recruited by adaptor proteins into focal

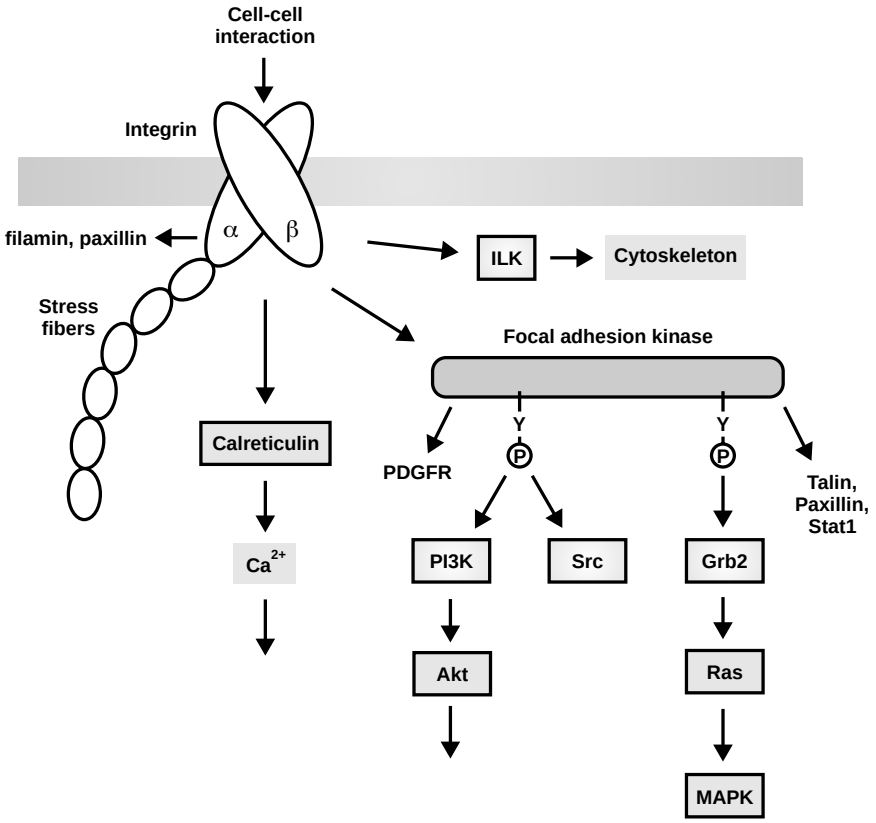


Fig. 11.13 Model of signal transduction via integrins. Activated integrins transmit signals from the extracellular matrix to the cytoskeleton and activate various intracellular signaling pathways. The diagram lists some of the signal proteins that have been shown to be involved in integrin signaling. The signal conduction involves - among others - the Ras/MAPK pathway and Ca^{2+}-signaling pathways, setting in motion a broad spectrum of subsequent reactions. At least two protein kinases participate in signal transduction, the integrin linked kinase (ILK) and the focal adhesion kinase (FAK). Furthermore proteins involved in reorganization of the cytoskeleton (paxillin, filamin) and in formation of stress fibers associate with activated integrins. Association of the Ca^{2+}-binding protein calreticulin with integrins is thought to link integrin signaling with Ca^{2+}-signaling. Activation and phosphorylation of FAK creates phospho-tyrosine residues functioning as attachment points for the recruitment of other signaling proteins like PI3-kinase, Src kinase and the adaptor protein Grb2.

adhesion points or other membrane microdomains containing integrins. The following proteins interact with integrins on the cytoplasmic side:

– microfilament or stress fiber components: α-actinin, filamin, talin, tensin
– adaptor proteins: paxillin
– membrane proteins: caveolin
– Ser/Thr-specific protein kinase: ILK, integrin-linked kinase

– nonreceptor tyrosine kinases: Syk, FAK (focal adhesion kinase)
– calreticulin, a Ca^{2+}-binding protein.

It is not known how the integrins and their downstream effectors cooperate structurally and functionally, how these multiprotein complexes are organized, and how ligand binding induces activation of the downstream effectors. It is assumed that ligand binding at the outer side induces a change in the mutual orientation of the α and β subunits, which then triggers binding and activation of the most proximate effector proteins. Most of the contacts to the downstream effectors of the integrins appear to be mediated by the integrin β-subunits. Many protein contacts have been mapped to the intracellular part of this polypeptide.

One of the first steps in integrin signaling is the activation of the nonreceptor tyrosine kinase Syk. Further steps include activation of Src kinase, FAK, ILK and recruitment of adaptor proteins like paxillin, Shc, Grb2 and others. Furthermore, integrin signaling is linked via actinin, talin, and other proteins to the actin cytoskeleton. Many of the integrin effector proteins are phosphorylated in the process of integrin signaling, and their phosphotyrosine residues serve as attachment points for SH2- or PTB-containing signaling proteins.

As a consequence of ligand binding and clustering of integrins, activation of different signaling pathways is initiated:
– activation of nonreceptor tyrosine kinases
– activation of Ser/Thr-specific proteins kinases such as the integrin-linked kinase (ILK)
– increase in Ca^{2+} concentration
– activation of the MAPK cascade
– increased formation of Ptd-Ins messenger substances.

Of the protein tyrosine kinases, the focal adhesion kinase (FAK, $p125^{FAK}$) plays an important role in integrin signal transduction (review: Schaller, 2001). The FAK protein undergoes autophosphorylation on integrin activation and has been shown to interact with other signaling proteins and with actin filaments. The phosphotyrosine residues of FAK serve as binding sites for SH2-containing signal molecules such as PI3 kinase, Src kinase and the adaptor protein Grb2 (Fig. 11.13). The FAK protein also has a specific binding domain for the adaptor protein paxillin and is found in a defined complex together with paxillin. Paxillin is a multifunctional docking protein that interacts with components of cytoskeleton structures (among other things), so recruitment of FAK to the cytoskeleton in the region of focal adhesion points seems possible via a paxillin-FAK interaction.

The integrin-mediated activation of MAPK pathways seems particularly important for integrin function, since this has an influence on transcription processes. The model of the mechanism of this linkage usually employs the Grb2/mSos complex and the Ras protein as a central switching station. Furthermore, there are also links between the integrins and the Rho/Rac GTPases. Here, it is interesting to note that the integrin-MAPK linkage can trigger the same biological events as growth factors that bind to transmembrane receptors.

In addition to directly generated signals, integrins can modulate signaling responses to external signaling molecules such as growth factors. The integrins influence the efficiency of signal transduction via receptor tyrosine kinases and the subsequent activation of the MAPK pathway. Integrin-mediated activation of the PDGF receptor has also been observed, which is independent of ligand binding at this receptor.

The link between cell anchoring and integrin activation seems to be of great biological importance. A link has been established between cell anchoring and regulation of the cell cycle machinery, although the exact nature of this link is not known. Integrin-mediated cell anchoring is also known to regulate a set of events impacting on apoptotic pathways. Along these lines, activation of PI3-kinase, Akt kinase (see Chapters 6 and 9), and influence on the Bcl-2 proteins and caspases (Chapter 15) have been established.

Reference

1 Arzt, E. (2001) gp130 cytokine signaling in the pituitary gland: a paradigm for cytokine-neuro-endocrine pathways. *J.Clin.Invest* , **108**, 1729–1733.

2 Becker, S., Groner, B. and Muller, C.W. (1998) Three-dimensional structure of the Stat3beta homodimer bound to DNA. *Nature*, **394**, 145–151.

3 Bromberg, J. and Chen, X. (2001) STAT proteins: signal tranducers and activators of transcription. *Methods Enzymol.*, **333**, 138–151.

4 Constantinescu, S.N., Ghaffari, S. and Lodish, H.F. (1999) The Erythropoietin Receptor: Structure, Activation and Intracellular Signal Transduction. *Trends Endocrinol.Metab*, **10**, 18–23.

5 Ellery, J.M. and Nicholls, P.J. (2002) Alternate signalling pathways from the interleukin-2 receptor. *Cytokine Growth Factor Rev.* , **13**, 27–40.

6 Frank, S.J. (2002) Receptor dimerization in GH and erythropoietin action–it takes two to tango, but how? *Endocrinology*, **143**, 2–10.

7 Garcia, K.C., Teyton, L. and Wilson, I.A. (1999) Structural basis of T cell recognition. *Annu.Rev.Immunol.*, **17**, 369–397.

8 Germain, R.N. (2001) The T cell receptor for antigen: signaling and ligand discrimination. *J.Biol.Chem.*, **276**, 35223–35226.

9 Horvath, C.M. and Darnell, J.E. (1997) The state of the STATs: recent developments in the study of signal transduction to the nucleus. *Curr.Opin.Cell Biol.*, **9**, 233–239.

10 Howe, A., Aplin, A.E., Alahari, S.K. and Juliano, R.L. (1998) Integrin signaling and cell growth control. *Curr.Opin.Cell Biol.*, **10**, 220–231.

11 Ihle, J.N., Thierfelder, W., Teglund, S., Stravapodis, D., Wang, D., Feng, J. and Parganas, E. (1998) Signaling by the cytokine receptor superfamily. *Ann.N.Y.Acad.Sci.*, **865**, 1–9.

12 Imada, K. and Leonard, W.J. (2000) The Jak-STAT pathway. *Mol.Immunol.*, **37**, 1–11.

13 Kerr, R., Stirling, D. and Ludlam, C.A. (2001) Interleukin 6 and haemostasis. *Br.J.Haematol.*, **115**, 3–12.

14 Krebs, D.L. and Hilton, D.J. (2001) SOCS proteins: negative regulators of cytokine signaling. *Stem Cells*, **19**, 378–387.

15 Livnah, O., Stura, E.A., Middleton, S.A., Johnson, D.L., Jolliffe, L.K. and Wilson, I.A. (1999) Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science*, **283**, 987–990.

16 Moutoussamy, S., Kelly, P.A. and Finidori, J. (1998) Growth-hormone-receptor and cytokine-receptor-family signaling. *Eur.J.Biochem.*, **255**, 1–11.

17 Qian, D. and Weiss, A. (1997) T cell antigen receptor signal transduction. *Curr.Opin.Cell Biol.*, **9**, 205–212.

18 Schaller, M.D., Otey, C.A., Hildebrand, J.D. and Parsons, J.T. (1995) Focal adhesion kinase and paxillin bind to peptides mimicking beta integrin cytoplasmic domains. *J.Cell Biol.*, **130**, 1181–1187.

19 Schwartz, M.A. (2001) Integrin signaling revisited. *Trends Cell Biol.*, **11**, 466–470.

12
Other Receptor Classes

The cell has other transmembrane receptors and signaling pathways that do not fit into the "classical" receptor types and signal mechanisms described in Chapters 5 – 11. The following signaling pathways certainly do not complete the list of intercellular and intracellular communication mechanisms in mammals, and it is to be expected that other classes of signaling pathways will be described in the future.

12.1
Receptors with Intrinsic Ser/Thr Kinase Activity: the TGFβ Receptor and the Smad Proteins

In addition to transmembrane receptors with intrinsic tyrosine kinase activity, the cell also contains transmembrane receptors with intrinsic Ser/Thr kinase activity. Examples of such a receptor type are the receptors for the *transforming growth factor β (TGFβ*) family of cytokines (review: Attisano and Wrana, 2002). Members of this family regulate a wide spectrum of cellular functions such as proliferation, apoptosis, differentiation, and cell migration. The TGF*β* family consists of secreted peptides of which the TGF*β*s, the activins, and the bone morphogenetic proteins (BMPs) are the best-known examples. These cytokines elicit their biological response through binding to heteromeric transmembrane receptors carrying an intrinsic serin/threonine kinase activity. On activation of the receptor, a signal chain is activated that leads signals to the transcription level and thus influences, e. g., the progress of the cell cycle. As an example, activation of the TGF*β* receptor by TGFß creates antimitogenic (i. e., inhibiting cell division) signals, which are manifested as increased production of inhibitors of cell cycle specific protein kinases (see Chapter 13).

12.1.1
TGFβ Receptor

The cytokines of the TGF*β* family are the activating ligands for the family of TGF*β* receptors. This comprises two subfamilies, the type I (T*β*R-I) and type II (T*β*R-II) receptors. These are transmembrane proteins composed of an extracellular ligand-binding domain, a single transmembrane element, and a cytoplasmic Ser/Thr kinase do-

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

main. The TGFβ receptor has a heterotetrameric structure in its activated form; it contains two copies of the TβR-I and TβR-II polypeptides. Several members are known for each of the TβR-I and TβR-II subfamily, and these are activated by specific cytokine ligands.

The mechanism of activation of the TGFβ receptor is shown schematically in Fig. 12.1 (review: Wrana and Attisano, 2000) for the TGFβ and activin-specific receptors.

The ligand TGFβ is a secreted polypeptide, which is cleaved from a large precursor protein and is found in dimeric form in the extracellular space. TGFβ first binds to a dimer of TβR-II, which carries a constitutively active serine/threonine kinase activity. TβR-I then binds to the complex of TGFβ and TβR-II, forming a hetero-oligomeric, ligand-bound receptor complex, in which phosphorylation of TβR-I by the kinase activity of TβR-II takes place. The phosphorylation of TβR-I occurs mainly in a Gly/Ser-rich domain (GS domain) located just N-terminal to the kinase domain. This phosphorylation activates TβR-I by creating a binding site for the downstream effectors, the Smad 2 and Smad 3 proteins, which are now phosphorylated and transmit the signal further. The ligand-induced formation of the hetero-oligomeric receptor and the subsequent phosphorylation of the GS domain convert the receptor into the active state from which the signal is transferred to effector molecules. This receptor type uses two subunits for signal transduction: TβR-II recognizes the signal, and TβR-I determines specificity for further signal conduction, in that it gives the signal to the downstream effector molecules.

A different binding and activation mode is used by the bone marrow morphogenetic factor (see review: Shah et al., 2001). In this case, the ligand BMP binds to and activates preformed heteromeric complexes of receptor types I and II.

From the activated TGFβ receptor, signals are directed to the transcription level. A set of proteins, known as *Smad proteins*, are involved in this signal conduction (see Fig. 12.1; reviews: Massague, 1998; Mehra and Wrana, 2002). As well as Smad proteins, G-proteins and the MAPK cascade are also involved in the downstream signaling of TGFβ family members.

12.1.2 Smad Proteins

At least nine different Smad proteins have been identified in higher organisms, and these may be divided into three functional classes:

- **Receptor-regulated Smad proteins** (R-Smads 1 – 3, 5, 8)
  The R-Smad proteins are effector molecules directly downstream of activated type I receptors. They bind to activated type I receptors and are phosphorylated and activated for further signal conduction by the receptor. Phosphorylation takes place on a C-terminal SSXS motif. Regulation of R-Smads by type I receptors is quite specific. R-Smads 2 and 3 are activated by TGFβ receptors and activin receptors, whereas R-Smads 1, 3 and 5 are regulated by the BMP receptors. Once activated

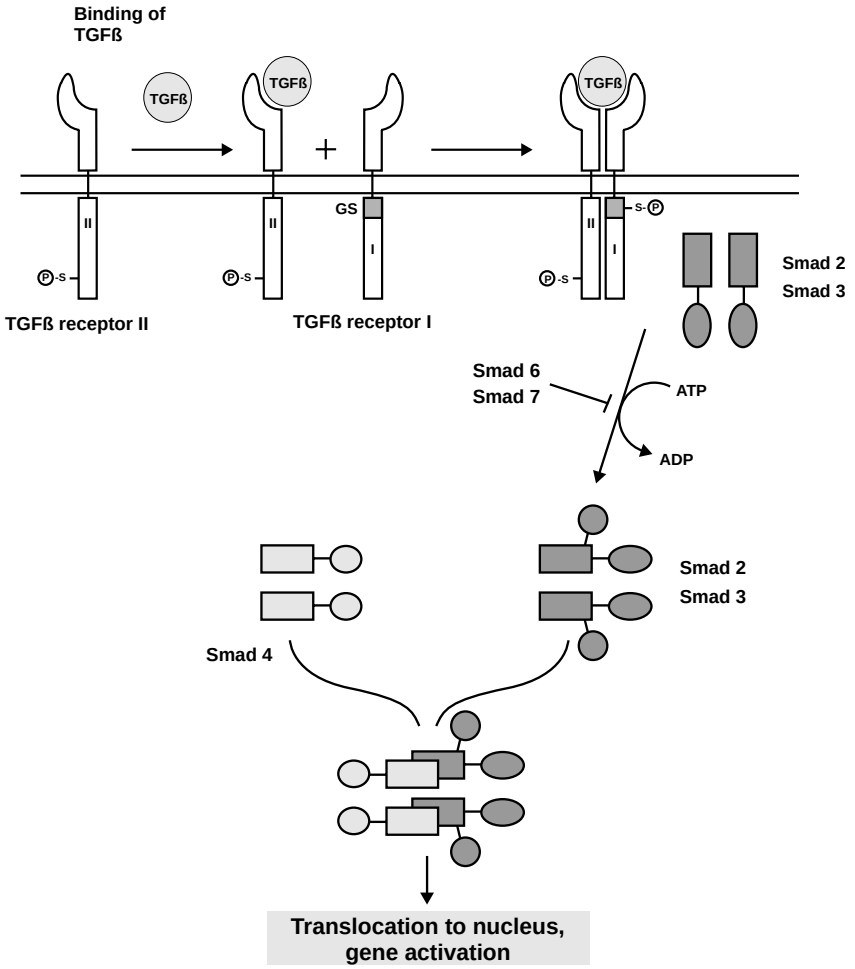


Fig. 12.1 Model of signal transduction via the TGFβ receptor. Signal transduction via TGFβ requires two TGFβ receptors, the TGFβ receptor I and the TGFβ receptor II. TGFβ first binds to TGFβ receptor II, which has a constitutive Ser/Thr-specific protein kinase activity. Binding of TGFβ to the TGFβ receptor II induces association of TGFβ receptor I with TGFβ receptor I; this phosphorylates the TGFβ receptor I on a glycine-serine rich domain (GS) and activates it for conduction of the signal further. The figure does not adress the correct oligomeric state of the activated receptor. The activated TGFβ receptor phosphorylates the pathway-restricted Smad proteins (Smad2 and Smad 3) which then associate with the common mediator Smad protein (Smad 4). The complex between Smad4 and phosphorylated Smad2/Smad3 translocates into the nucleus where it binds to specific DNA elements of target genes and activates transcription. The inhibitory Smads (Smad 6, Smad 7) can also bind to the TGFβ receptor and may prevent phosphorylation of Smad 2/Smad3 by the TGFβ receptor (According to Heldin et al., 1997). TGFβ: transforming growth factor type β.

by phosphorylation, the R-Smads dissociate from the receptor and are found in trimeric or higher order complexes in the cytosol.
The activity of the R-Smad 2 is also regulated by another class of proteins, named SARA, which control the cytoplasmic transport and the membrane association of Smad 2.

- **Common mediator Smads, co-Smads**
  The mode of action of Smad 4 clearly differs from that of the other members of the Smad family. Smad 4 binds to phosphorylated R-Smads and forms trimeric complexes composed of two R-Smad molecules and one Smad 4 molecule. These complexes translocate to the nucleus, where they bind to related DNA elements and activate the transcription of target genes. The mechanism of transcription regulation by Smads is complex and includes both positive and negative influences. Generally, Smad-dependent regulation of transcription requires the interaction with other transcription factors, such as members of the FoxH1 family of forkhead transcription factors, the Vitamin D receptor and the c-Jun transcription factor, among others (review: Attisano et al., 2001). Futhermore, Smads can interact with coactivators and corepressors of transcription and thereby recruit, e. g., histone acetylase activity or histone deacetylase activity to chromatin.
- **Inhibitory Smads (I-Smads)**
  Smad 6 and Smad 7 are I-Smads, i. e. they function as inhibitors of TGFβ signaling. Several mechanisms for this have been identified. They can bind to type I receptors and interfere with phosphorylation of the pathway-restricted Smads. Furthermore, the I-Smads can interfere with the interaction between Smad 3 and Smad 4, and can recruit an E3 ubiquitin ligase named Smurf to the receptor, which induces ubiquitination and proteasomal degradation of the receptor complex. These properties of the I-Smad proteins are used to modulate and weaken TGFβ receptor signaling. Such transmodulation has been described for the interferon γ pathway. Activation of the IFNγ pathway leads to increased transcription of Smad 7, which diminishes signal conduction in the TGFβ pathway (Ulloa et al., 1999).

Comparison of Smad proteins has revealed the presence of conserved N- and C-terminal domains named MH1 and MH2. These domains function by controlling protein-DNA and protein-protein interactions.

DNA binding of Smad 3 takes place via a DNA-binding motif within MH1 which differs substantially from classical DNA-binding motifs. The DNA-binding domain of Smad 3 contains a β-sheet motif, which positions itself in the large groove of the DNA (Shi et al., 1998). The various R-Smads differ only slightly in their DNA-binding specificity, which is in line with the observation that the R-Smads require cooperation with other, sequence-specific transcription factors for transcription activation.

The TGFβ/Smad signaling pathway is controlled at various levels and is subjected to crosstalk with other signaling pathways (Fig. 12.2). The main regulatory influences are listed below (review: Kloos et al., 2002).

– binding of I-Smads
– targeted degradation of the TGFβ receptor: recruitment of a ubiquitin E3 ligase by I-Smad 7

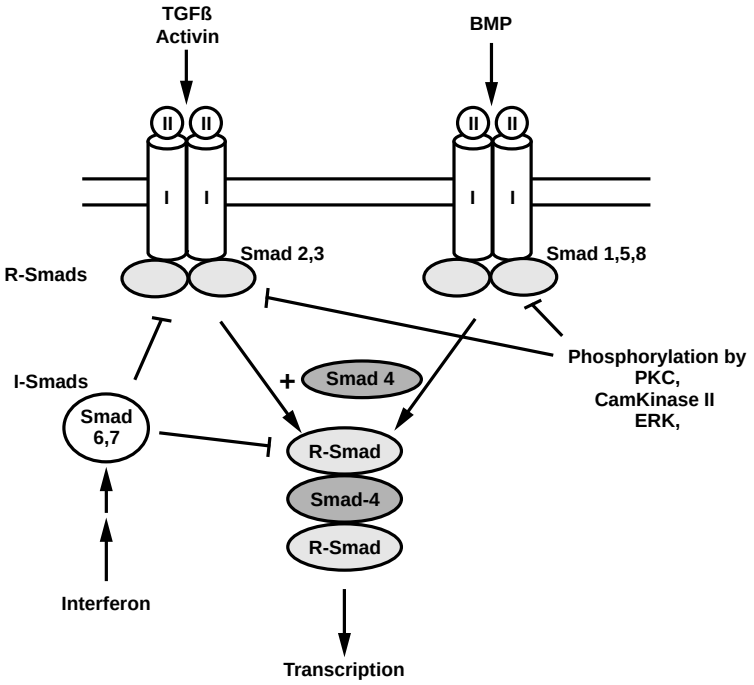


Fig 12.2 Regulation of signaling by the TGFβ superfamily receptors. The extracellular signaling molecules TGFß, activin and bone morphogenetic factor (BMP) each activate a distinct receptor composed of type I and type II subunits. TGFß and activin receptors activate the Co-Smads Smad2 or Smad3, the BMP receptor uses Smad 1, 5 or 8 as Co-Smads for further signal transmission. Smad6 and Smad7 inhibit signaling by interfering with complex formation between the R-Smads and the Co-Smad or by inhibiting phosphorylation of the R-Smads by the activated receptor. R-Smad phosphorylation can be negatively regulated by protein kinase C (PKC), CamKinase II and the extracellur regulated kinase, ERK.

– phosphorylation and inhibition of R-Smad function via ERK2 of the MAPK pathway
– phosphorylation and inhibition of R-Smads by CaM kinase II
– phosphorylation and inhibition of R-Smads by protein kinase Cγ.

In total, TGFβ-Smad signal conduction has distinct similarities to signal conduction in the Jak-Stat pathway (see Section 11.2.2). In both pathways, cytosolic transcription factors are activated by phosphorylation and are translocated in oligomeric complexes to the nucleus and the DNA. Common to both pathways is the short distance from the extracellular signal to the transcription level. The existence of multiple regulatory influences on Smad signaling suggests that Smads function as signal integrators of multiple signaling pathways that allow fine tuning and output modification of other signaling pathways. At the transcriptional level, the response to TGFβ family member ligands appears to be dependent upon what other signals are being received during transcriptional activation, allowing a variety of extracellular stimuli to control the transcriptional activity.

12.2 Receptor Regulation by Intramembrane Proteolysis

A surprising new concept of transmembrane signaling has emerged from studies on the processing of transmembrane proteins, including the Notch protein, the sterol regulatory element-binding protein (SREBP), the Alzheimer precursor protein (APP), the receptor tyrosine kinase ErbB4 and the cell adhesion molecule E-cadherin.. These proteins appear to be processed and activated by a process named regulated intramembrane proteolysis (RIP, reviews: Brown et al., 2000; Urban and Freeman, 2002). The pathway uses several proteolytic steps to release a regulatory protein fragment from the intramembrane portion of a transmembrane protein. Often the released fragments function as transcriptional regulators and specific changes in gene activity are observed as a consequence of the activation of the RIP pathway. This activation is triggered in most cases by ligand binding to the transmembrane protein.

The principles of this type of signaling will be explained using the example of Notch signaling. The Notch-signaling pathway mediates cell-cell contacts used to control the cell fate of neighboring cells in neuronal differentiation processes. The Notch proteins are ligand-activated transmembrane receptors with the transmembrane proteins of neighboring cells as ligands. These ligands belong to the family of delta/serrate proteins. On binding a ligand during a cell-cell interaction, a direct signal is transmitted to the transcription level, and various target cells are activated.

The unstimulated Notch receptor is found as a heterodimeric transmembrane protein consisting of a large ectodomain that is noncovalently associated with the Notch transmembrane intracellular domain (Fig.12.3). The two polypeptides are generated by cleavage from a large precursor protein, and this cleavage is catalyzed in the Golgi network by a protease named furin. Ligand binding at the cell surface results in specific proteolysis of the heterodimer at a position 12 – 13 amino acids outside the membrane and removal of the extracellular portion of the heterodimer. A protease named TACE catalyzes this second proteolytic step. The cytosolic fragment which still has the transmembrane element and a short extracellular part attached is then further processed by cleavage near the cytosolic end of the membrane-spanning segment. As a result of this third proteolytic step, the cytosolic domain of Notch is released from the membrane and is translocated into the nucleus, where it activates transcription as part of a complex with a family of transcription factors known as CSL. CSL proteins (e. g., CBF1 in mammals) can bind to specific sites on the DNA, activating the transcription of basic helix-loop-helix proteins (see Section 1.2.1), which in turn affect the transcription of downstream target genes.

The protease responsible for the critical third proteolytic step is contained in a large protease complex named γ-secretase (review: Fortini, 2002). Proteins belonging to the presenilin family of proteases have been identified as the catalytic component of the γ-secretase complex. The presenilins are transmembrane proteases with six transmembrane elements that use aspartate residues to cleave substrates. They have also been implicated in the processing of the Alzheimer precursor protein (APP), and mutations of the presenilins appear to contribute to the pathogenesis of Alzheimer's disease. In

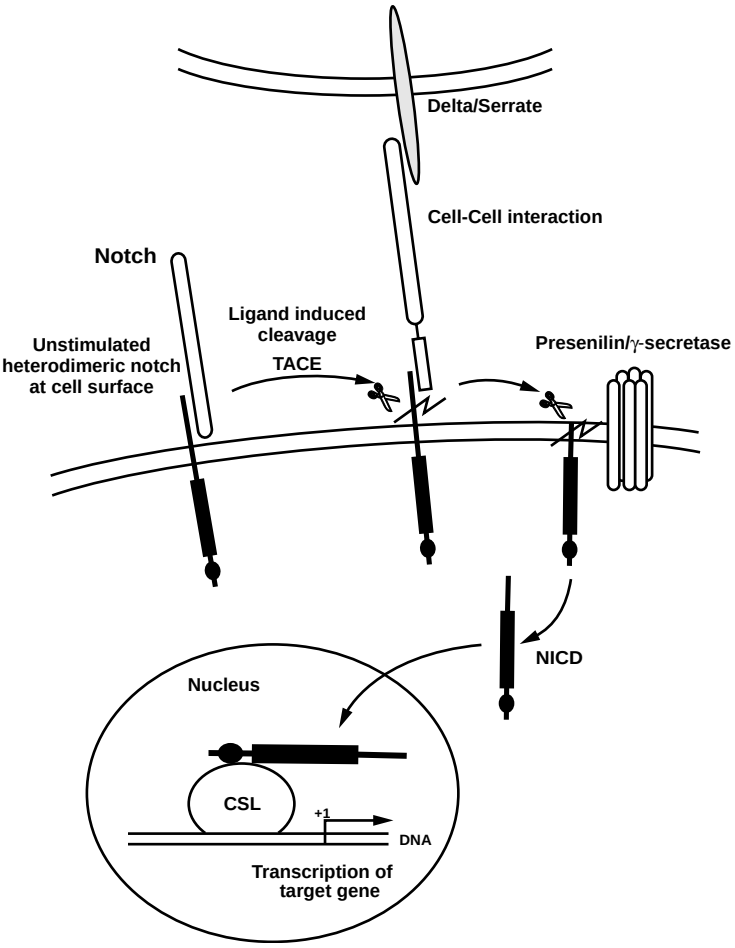


Fig. 12.3 Model of Notch signaling. Notch receptors are found as heterodimers at the cell surface. In the course of cell-cell interactions, binding of Notch ligands (e. g. delta, serrate proteins) to Notch induces cleavage by the TACE proteases at a site close to the transmembrane segment. Presenilin/γ-secretase then mediates cleavage at a site located within the transmembrane segment to release the Notch intracellular domain (NICD) from the membrane. NICD then enters the nucleus where it forms a complex with DNA-bound CSL protein to activate the transcription of target genes, including genes for basic helix-loop helix transcription factors.

addition to the presenilins, three other families of transmembrane proteases are known (review: Urban and Freeman, 2002)

The Notch signaling pathway is thus a further example of signal conduction in which the signal is directed by the shortest pathway from the cell membrane into the nucleus.

12.3 Signal Transduction via the Two-Component Pathway

The two-component pathway was originally discovered in bacteria. It was only recently recognized that this kind of signal transduction is also used in eucaryotes. Bacteria possess signal systems which they use to react to changes in N availability, osmolarity, and chemotactic substances. The signaling pathway responsible for this regulation is known as the *two-component pathway*, because of the involvement of two conserved functional and structural domains.

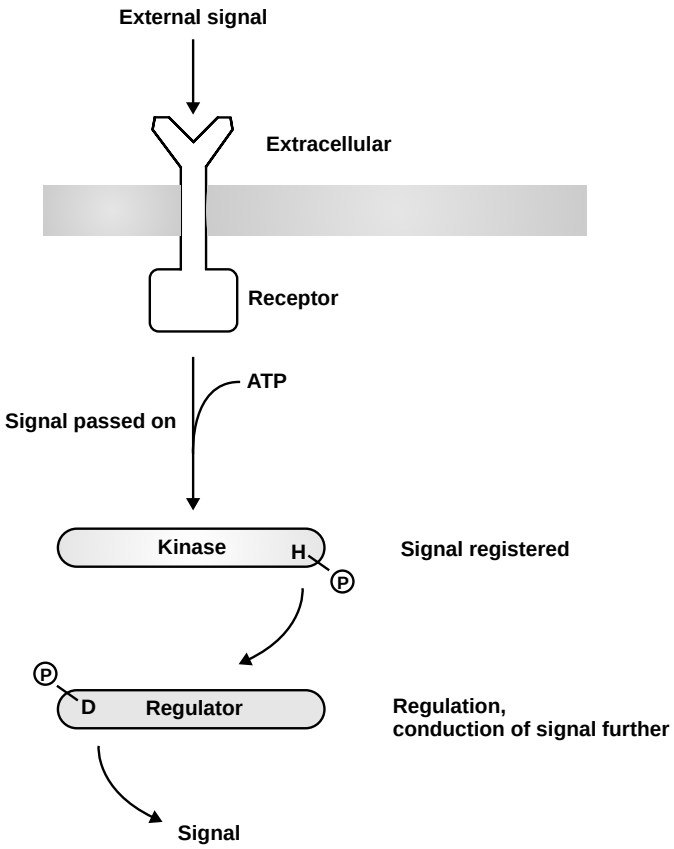


Fig. 12.4 Principle of the two-component pathway. The figure shows the principal steps of the two-component pathway in bacterial systems. An extracellular signal (change in osmolarity, N availability, etc.) is registered by a receptor. An interaction takes place with the first component, the "sensor kinase", which undergoes autophosphorylation at a His residue (H). The phosphate residue is transferred to the carboxyl side chain of an Asp residue (D) of the reaction regulator. Phosphorylation of the second component activates this for further signal conduction. The "sensor kinase" may also be localized in the cytoplasmic domain of the receptor.

The two-component pathway is characterized by two functional elements. A *histidine-specific protein kinase* functions as a "*sensor*" that registers an external signal and passes this on to a downstream *response regulator*. The latter is activated by phosphorylation during the process of signal transduction, triggering other reactions in the cell (Fig. 12.4).

The composition of the two-component pathway is very variable. The nature of the external signal and the reactions triggered in the cell may be very diverse. The sensor kinase may be a part of the receptor that registers the signal, or it may be on a polypeptide chain separate from the receptor. Furthermore, there are different mechanisms of coupling of the main functions of the two-component pathway. The sensor and reaction regulator may be on a single polypeptide or they may be on separate proteins. In addition, the proteins involved may be membrane proteins.

In the two-component pathway, external signals are registered by a His-specific protein kinase. The His-specific protein kinase is activated as a reaction to reception of an external signal (e. g., binding of a ligand) and is followed by autophosphorylation of a conserved *His residue* using ATP as a phosphate donor. The presence of a His kinase domain in a protein of the two-component pathway can often be diagnosed on the basis of the primary sequence, because of the comprehensive sequence information available (more than 50 bacterial systems are known). The signal is passed on to the *regulator* from the phosphohistidine of the activated kinase, with the phosphate residue being transferred to a conserved *Asp residue* in the regulator. The Asp phosphorylation of the regulator controls its activity in signal transduction. Downstream proteins may be activated or the signal may be directly converted into regulation of transcription. A DNA-binding domain is often found in the regulator, so that DNA-binding capacity of the regulator is controlled by Asp phosphorylation.

The two-component system is a signaling pathway of great importance in bacteria. Similar proteins and signaling pathways can also be identified in plant cells and in yeast, based on sequence homology with the bacterial proteins. It is expected that signaling pathways using the principle of the two-component system will also be found in mammals.

Fig. 12.5 shows a model for signal transduction in the two-component system of yeast (review: Wurgler-Murphy and Saito, 1997). The SLN1 protein of yeast is a composite sensor-regulator system in which the His kinase domain and the regulator domain are located on the same protein. Following reception of an extracellular signal (e. g., low osmolarity), autophosphorylation takes place on His and Asp residues of the SLN1 protein. In the process of signal transduction, a phosphate residue is transferred from a His phosphate of the SLN1 protein to an Asp residue on a downstream signal molecule SSK1. In the unphosphorylated form, SSK1 directs a signal to the MAP kinase pathway of yeast. Asp phosphorylation of SSK1 stops this signal conduction. The system is an example of a two-component system in which Asp phosphorylation takes place on both the response regulator and on the substrate protein.

A particular feature of the two-component system is the use of phosphohistidine and phosphoaspartate groupings (see Fig. 7.1) as elements of signal transduction. The phosphonamide grouping in phosphohistidine is unstable in aqueous medium and has only a short half-life. The duration of the signal transduction is thus limited to a large extent by the chemical instability of phosphohistidine.

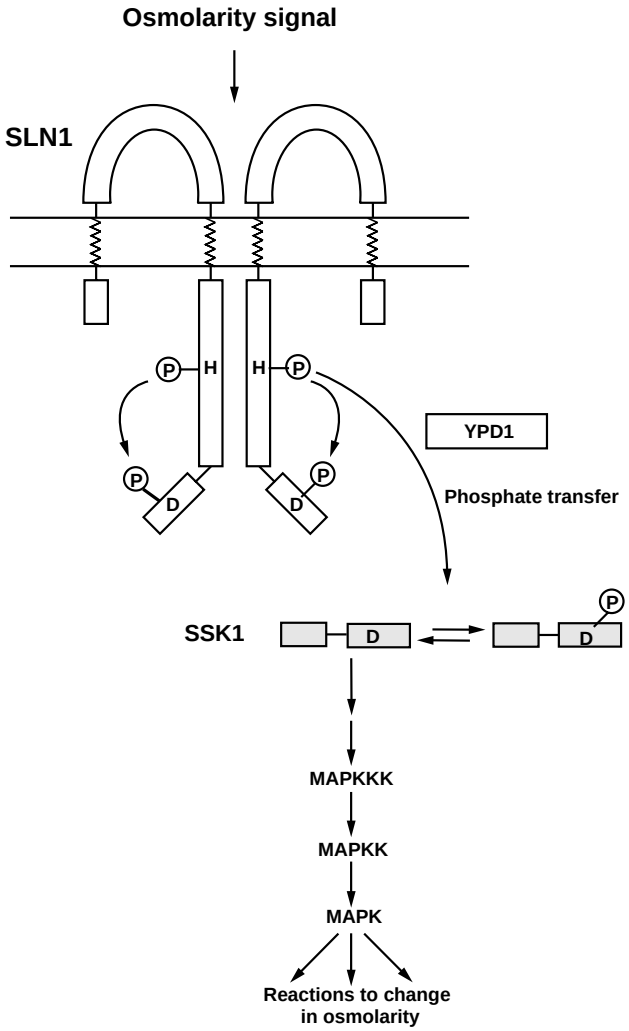


Fig. 12.5 Example of a two-component pathway in *S. cerevisiae*. Model of signal transduction via the SLN1 protein. The SLN1 protein is a transmembrane protein with two transmembrane elements, which is assumed to exist as a dimer. The sensor domain and the regulator domain are localized on the same protein chain in the SLN1 protein. The SLN1 protein is activated by an extracellular signal (e. g., decrease in osmolarity). Autophosphorylation takes place on His (H) in the sensor domain and on Asp (D) in the regulator domain. A phosphate transfer takes place from the phosphohistidine to the effector protein SSK1. Another protein, named YPD1 participates in this transfer. In the unphosphorylated form, SSK1 activates a MAPK pathway, which contains the protein kinase HOG1 as a MAPK element. Various cellular reactions are triggered by HOG1. If SSK1 is phosphorylated in the course of activation of the two-component pathway, stimulation of the MAPK pathway is stopped.

The switch for weakening and activation of the signal is already built in to the system. If there is no regulator protein available to receive the signal, the signal dies away because of the chemical instability of His phosphate. This is similar for the Asp phosphate grouping of the regulator. Here, there are specific phosphatases that can cleave the phosphate off again, but the intrinsic instability of the aspartyl phosphate grouping can be large and is determined by the microenvironment of the protein. Thus, Asp phosphate in the bacterial Che-Y protein has a half-life of only a few seconds. In contrast, the OmpR protein has an Asp phosphate with a half-life of ca. 1 h.

The high intrinsic instability of the groupings involved in this chemical signal transduction makes the two-component system particularly suitable for rapid, repeated signal transductions. It is a relatively simple system in which a signal may be rapidly conducted further and rapidly switched off.

Reference

1 Attisano, L., Silvestri, C., Izzi, L. and Labbe, E. (2001) The transcriptional role of Smads and FAST (FoxH1) in TGFbeta and activin signalling. *Mol.Cell Endocrinol.*, **180**, 3–11.

2 Attisano, L. and Wrana, J.L. (2002) Signal transduction by the TGF-beta superfamily. *Science*, **296**, 1646–1647.

3 Brown, M.S., Ye, J., Rawson, R.B. and Goldstein, J.L. (2000) Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. *Cell*, **100**, 391–398.

4 Fortini, M.E. (2002) Gamma-secretase-mediated proteolysis in cell-surface-receptor signalling. *Nat.Rev.Mol.Cell Biol.*, **3**, 673–684.

5 Kloos, D.U., Choi, C. and Wingender, E. (2002) The TGF-beta–Smad network: introducing bioinformatic tools. *Trends Genet.*, **18**, 96–103.

6 Massague, J. (1998) TGF-beta signal transduction. *Annu.Rev.Biochem.*, **67**, 753–791.

7 Mehra, A. and Wrana, J.L. (2002) TGF-beta and the Smad signal transduction pathway. *Biochem.Cell Biol.*, **80**, 605–622.

8 Shah, P.K., Buslje, C.M. and Sowdhamini, R. (2001) Structural determinants of binding and specificity in transforming growth factor-receptor interactions. *Proteins*, **45**, 408–420.

9 Shi, Y., Wang, Y.F., Jayaraman, L., Yang, H., Massague, J. and Pavletich, N.P. (1998) Crystal structure of a Smad MH1 domain bound to DNA: insights on DNA binding in TGF-beta signaling. *Cell*, **94**, 585–594.

10 Ulloa, L., Doody, J. and Massague, J. (1999) Inhibition of transforming growth factor-beta/SMAD signalling by the interferon-gamma/STAT pathway. *Nature*, **397**, 710–713.

11 Urban, S. and Freeman, M. (2002) Intramembrane proteolysis controls diverse signalling pathways throughout evolution. *Curr.Opin.Genet.Dev.*, **12**, 512–518.

12 Wrana, J.L. and Attisano, L. (2000) The Smad pathway. *Cytokine Growth Factor Rev.*, **11**, 5–13.

13 Wurgler-Murphy, S.M. and Saito, H. (1997) Two-component signal transducers and MAPK cascades. *Trends Biochem.Sci.*, **22**, 172–176.

13
Regulation of the Cell Cycle

13.1
Overview of the Cell Cycle

Eucaryotic cells execute their reproduction in a cyclic process, in which at least two phases, *S phase* and *M phase*, can be differentiated on the basis of biochemical and morphological features. The biochemical characteristic of S (synthesis) phase is the replication of nuclear DNA and thus doubling of the genetic information. In M (mitosis) phase, division of the chromosomes between the daughter cells is prepared and carried out.

In most cell types, two further phases can be distinguished, G_1 and G_2 phase. G_1 phase covers the period between M phase and S phase while G_2 phase covers the period between S phase and M phase. From G_1 phase, the cell may transfer into a quiescent state known as G_0 phase. Appropriate signals (e. g., addition of growth factors) can induce the cell to return from G_0 into G_1 phase and proceed with the cell cycle.

The cyclical sequence of G_1, S, G_2 and M phases describes a standard cell cycle (Fig. 13.1). Rapidly dividing cells in mammals require 12 – 24 h for completion of a cell cycle. In some cell types, such as early embryonal cells, the period between the S and the M phases is reduced to the extent that discrete G_1 and G_2 phases cannot be identified. The duration of the cell cycle is then only 8 – 60 min.

Morphologically, cell division is only visible in M phase. Under the light microscope, condensation, alignment and segregation of the chromosomes and cell division itself may be observed during M phase. In addition, different mitotic phases can be distinguished, as shown in Fig. 13.2.

13.1.1
Principles of Cell Cycle Control

The different phases of the cell cycle include a number of highly ordered processes that ultimately lead to duplication of the cell. The various cell cycle events are highly coordinated to occur in a defined order and with an exact timing, requiring precise control mechanisms.

The ordered sequence of cell cycle events is ensured by different control loops that have an inhibitory or promoting effect on the progress of the cell cycle. These are

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

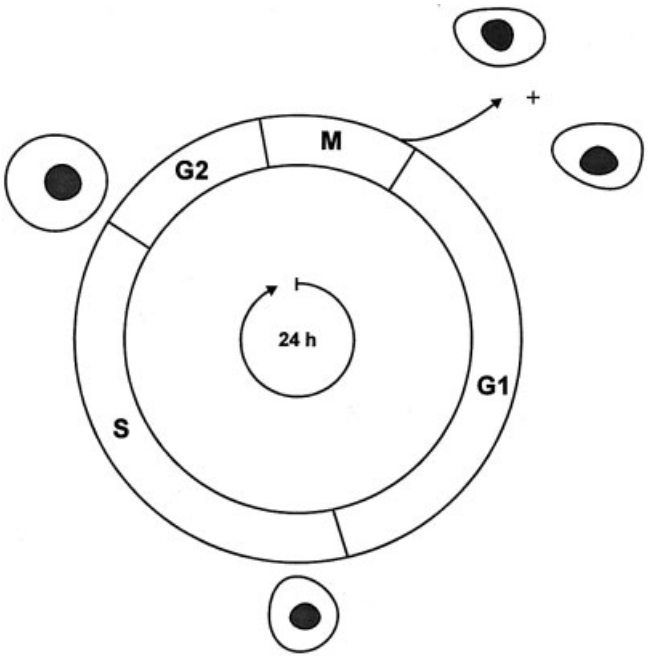


Fig. 13.1 The four phases of a typical cell cycle of a eucaryotic cell. G_1-, S- and G_2 phases form the interphase, whilst the cell grows continuously. Cell division occurs in M phase. New synthesis of DNA is limited to S phase. G_1 phase includes the period between M phase and S phase; G_2 phase lies between S phase and M phase. The duration shown of 24 h is only approximate. There are cells with cell cycles of shorter or longer duration.

monitoring mechanisms that register the completion of important cell cycle events (e. g., complete DNA synthesis) and allow the transition to the next event (e. g., entry into mitosis) to occur.

The control systems of the cell cycle ensure that the various phases are executed completely and in the correct sequence. Entry into a new phase can only take place when the preceding phase has been completed. In addition, the system allows coupling of processes that are not adjacent in the cell cycle sequence. Thus, there are close mutual regulation mechanisms between M and S phase. Examples of the various control mechanisms are shown in Fig. 13.3.

Many of the control mechanisms of the cell cycle are of an *intrinsic* nature and are constitutive, i. e., they are operational in every cell cycle and ensure the ordering of the individual steps. However, other control mechanisms exist that are not active in every cell cycle; these are only induced when defects are detected in central cell cycle events. These control mechanisms are known as *checkpoints*. An example of a checkpoint that is only activated when required is the DNA damage checkpoint (see Section 13.8). This is a biochemical pathway that detects DNA damage and creates a signal that slows cell cycle progression or arrests cells in the G_1, S or G_2 phase.

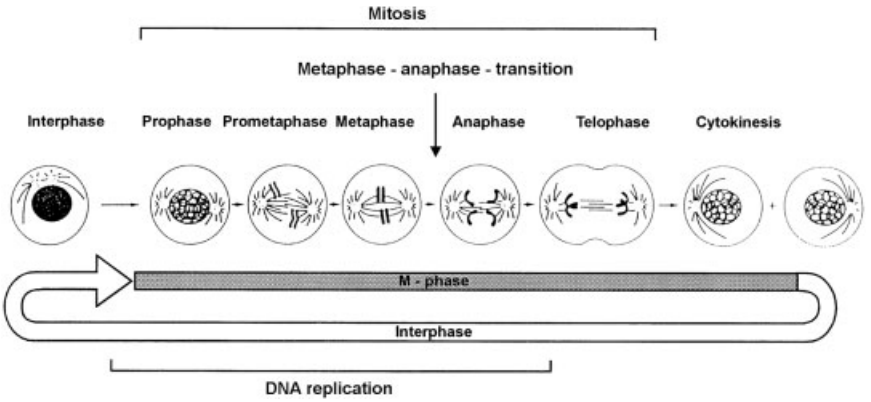


Fig. 13.2 Cytologic features of M phase. M phase is divided into the mitotic phases shown, based on characteristic cytologic features. The transition from metaphase to anaphase is an important control point. Cells may stop and pause before this control point. If the control point is crossed, M phase is concluded with cell division.

In addition to the built-in protection and control mechanisms, the cell is also subject to a number of *external* controls, which ensure that cell division occurs in balance with the overall development of the organism and with external growth conditions. This is a kind of social control of cell division that regulates the progress of the cell cycle, with the help of circulating signal molecules or via cell-cell interactions.

A biochemical system is at the center of the cell cycle, of which the most important players are Ser/Thr-specific protein kinases and regulatory proteins associated with these. The activity of this central cell cycle apparatus regulates processes downstream that help to carry out the many phase-specific biochemical reactions of the cell cycle in a defined order.

Controlling influences on the progress of the cell cycle are effected from various levels that are linked to one another (Fig. 13.3).

13.1.2 Intrinsic Control Mechanisms

Intrinsic control mechanisms ensure that the cycle is executed completely, so that, following cell division, both daughter cells are equipped with the same genetic information as far as possible. Of the many control mechanisms, the following are highlighted:

- **Coupling of mitosis to a completed S phase**
  Mitosis is only initiated when the DNA has been completely replicated during S phase. Mechanisms must exist that register completion of S phase and couple this to entry into M phase.

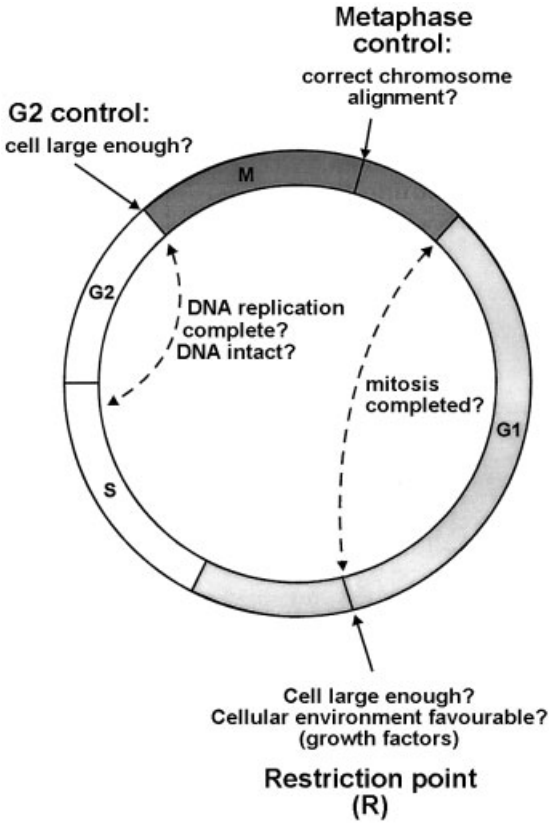


Fig. 13.3 Control points of the cell cycle: external and internal control mechanisms. Important control points of the cell cycle lie at the end of G_2 phase (G_2/M transition), in mitosis (metaphase/anaphase transition) and in G_1 phase (restriction point). The internal controls are shown as broken arrows and the external controls are shown as solid arrows.

- **Coupling of S phase and mitosis**
  Another control mechanism ensures that entry into S phase is only possible if preceded by mitosis. If the cell was able, during G_2 phase, to enter a new S phase without mitosis taking place, this would lead to unprogrammed multiplication of the chromosome set and thus to polyploidy. For S-phase control, see Section 13.5.
- **Coupling of cell size and progress in G_1 phase**
  A further control mechanism, which is also intrinsic, tests whether the cells in G_1 phase are large enough to initiate another round of cell division. The daughter cells produced by cell division must reach a critical size in the course of G_1 phase before S phase can commence.
- **DNA damage and the course of the cell cycle**
  The cellular genome is continually subject to damaging influences that may originate within the cell or externally (see also Section 14.1.4). DNA damage must be repaired with the help of repair enzymes. Nonrepaired DNA damage leads to miscoding in DNA replication and thus to mutations. For this reason, the cell has control mechanisms that register DNA damage and may possibly halt the cell cycle. Thus, time is gained for repair of DNA damage. However, the cell may also enter a resting state and may possibly initiate programmed cell death.

13.1.3
External Control Mechanisms

Growth Conditions

Cell division activities are controlled to a high degree by externally determined growth conditions such as nutrient supply. A cell may stop cell division if the physiological conditions are unfavorable.

Mitogenic Signals During Cell–Cell Communication

Within the bounds of intercellular communication, mitogenic signals in the form of growth factors are produced in the organism. These bind to specific receptors on the target cell and initiate signal chains that influence the progress of the cell cycle

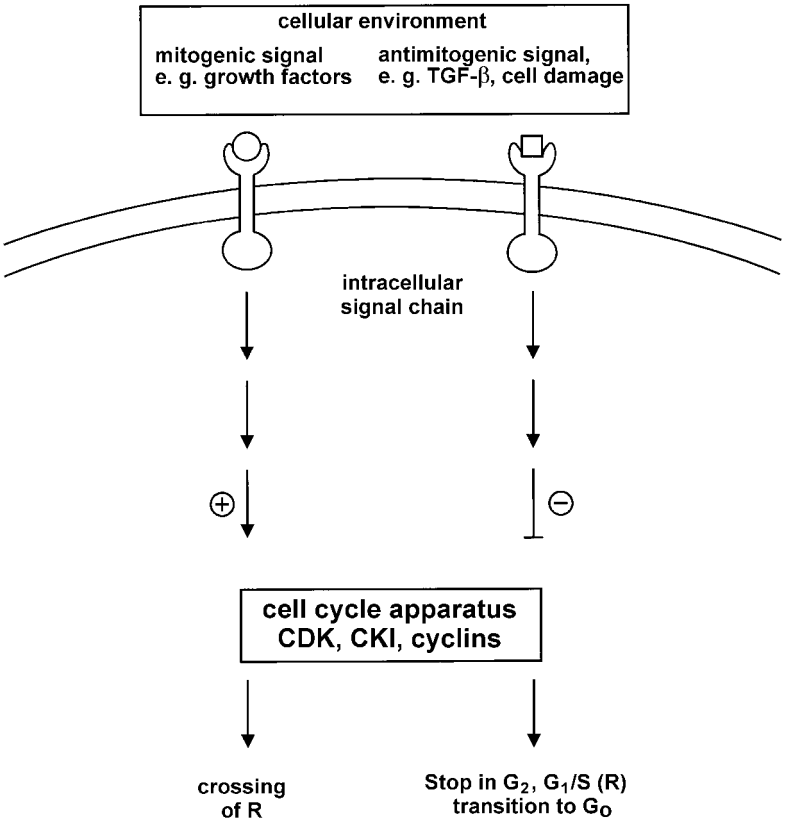


Fig. 13.4 Mitogenic and antimitogenic signals in control of the cell cycle. The cellular environment may emit mitogenic or antimitogenic signals. Mitogenic signals (e. g., growth factors) promote passage through the cell cycle; antimitogenic signals (e. g., TGFβ) lead to a halt in the cell cycle. In both cases, the extracellular signal is registered by transmembrane receptors and is passed on to the cell cycle apparatus via an intracellular signal chain. TGFβ: transforming growth factor β; CDK: cyclin-dependent protein kinase; CKI: inhibitor of CDK; R: restriction point.

(Fig. 13.4). These mitogenic signals control cell division activity and adjust it to the overall development of the organism.

Antimitogenic Signals during Cell–Cell Communication

In addition to growth-promoting signals, growth-inhibiting antimitogenic signals may also take effect in the organism. These lead to a halt in the cell cycle and may lead to transition of the cell into G_0 phase. Lack of mitogenic signals can have the same effect on the progress of the cell cycle.

13.1.4 Critical Cell Cycle Events and Cell Cycle Transitions

The cell cycle contains critical events where the cell switches from one state of biochemical activity to another, in an irreversible manner. These events are called cell cycle transitions. Often, activating and inhibitory signals are received and transmitted at these cell cycle transitions (Fig. 13.3).

An example of an important cell cycle transition is the restriction point R, which occurs in late G_1 phase. Crossing the restriction point is an important decision for further progression in the cell cycle and for entry into S phase. At this point, the cell switches from a growth-factor-dependent state to a growth-factor-independent state.

The restriction point is crossed, for example, when the cell is large enough for division and when enough external activating signals (growth factor signals) are present during G_1 phase, so that the cell is sufficiently prepared for the events following in S phase. On crossing the restriction point, the cell continues in the cell cycle automatically and there is no need for further activating signals for entry into S phase and continued progress in the cell cycle. If the conditions for crossing the restriction point are not fulfilled, progress through the cell cycle slows down or the cell stops in the cycle until the requirements are fulfilled.

Other important cell cycle transitions are entry into S phase and the G_2/M transition. At the G_2/M transition, it is registered whether S phase has been completely executed, and the integrity of the DNA is examined at a DNA damage checkpoint. There are other important cell cycle transitions in M phase between metaphase and anaphase. At this point, an irreversible decision is made for the progress of mitosis: if the spindle apparatus is correctly formed and the sister chromatids are correctly aligned, the cell cycle may proceed.

13.2 Key Elements of the Cell Cycle Apparatus

Investigations of the cell cycle of yeast have helped to identify the key elements of cell cycle regulation and to understand the underlying principles. The results obtained are mostly applicable to higher eucaryotes; here too, a plausible picture has been obtained

of how the cell cycle is maintained and regulated. However, many aspects of this picture are incomplete, and we are a long way from being able to understand and biochemically describe the coordinated progress and regulatory aspects of the cell cycle.

Two processes are central to cell cycle regulation:

– oscillating changes in the activity of the cell cycle machinery, with protein kinases as the most important component
– specific proteolysis of cell cycle regulators (see Section 13.3).

The activity of the cell cycle machinery is controlled by the following proteins in particular:

– cyclin-dependent protein kinases (CDKs)
– cyclins
– inhibitors of cyclin-dependent protein kinases (CKIs).

An oscillating system is formed by the interplay of the three protein classes, and the activity of this system makes up the specific biochemical functions of the individual phases of the cycle. The activity of the cyclin-dependent protein kinases (CDKs) is central to the oscillating system. These create a signal that initiates downstream biochemical processes and thus determines the individual phases of the cycle. CDK activity is also the starting point for intrinsic and external control mechanisms.

13.2.1 Cyclin-dependent Protein Kinases, CDKs

The CDKs are proteins of 34 – 40 kDa with Ser/Thr-specific protein kinase activity. The CDKs must associate with the corresponding cyclin (or cyclin-like proteins) to be active (review: Obaya and Sedivy, 2002).

Active cyclin-dependent protein kinases are thus heterodimers in which the CDK subunit carries the catalytic activity and the other subunit, the cyclin, performs an activating and specificity-determining function. In addition to association of the cyclin, most CDKs require phosphorylation in the activation segment (T loop or activation segment, see Section 7.2.1) for full activation.

In the fission yeast *Schizosaccharomyces pombe*, the oscillator function is performed by only one CDK subunit, the CDC2-CDK (also known as p^{34cdc2}); in the budding yeast *Saccharomyces cerevisiae*, this is CDC28-CDK ($p^{34cdc28}$). In mammals, there are at least ten different CDKs, numbered from CDK1 to CDK10 (see Table 13.1).

CDKs involved in Cell Cycle Regulation

Progress of the cell cycle appears to be mainly controlled by CDK1 (also known as CDC2) and CDKs 2, 4 and 6.

Tab. 13.1 CDKs, cyclins and CKIs in mammals and in the yeast *S. cerevisiae*.

Cyclin	CDK	CKI	Phase/function
Mammals			
A1	CDC2, CDK2	p21, p27, p57	Meiosis
A2	CDC2, CDK2	p21, p27, p57	S, GS, M
B1, B2, B3	CDC2	?	M
C	CDK8	?	Transcription repression
D1, D2, D3	CDK2, 4, 5, 6	p15, p16, p18, p19, p21, p27	G1, Restriction point
E	CDK2	p21, p27, p57	G1/S
F	?	?	G2, binding to cyclin B
H	CDK7	?	CDC2 phosphorylation CTD phosphorylation
T	CDK9	?	CTD phosphorylation, HIV-TAT target
p35, p39	CDK5		Phosphorylation of structural proteins
	CDK10		Phosphorylation of transcription factors
S. cerevisiae			
CIn1, CIn2, CIn3	Cdc28	Sic1	G1
CIb1, CIb2	Cdc28	?	M
CIb3, CIb4	Cdc28	?	G2
CIb5, CIb6	Cdc28	Sic1	S

Other CDKs

The CDKs 5 and 7-10, together with the corresponding cyclins, are not directly involved in cell cycle control and perform other specific tasks, e. g., in transcription regulation.

- **CDK5**
  This CDK requires the cyclin-related proteins p35 or p39 as activating subunits. CDK5 (review: Dhavan and Tsai, 2001) regulates the architecture of cells of the nervous system by phosphorylation of structural proteins like dynamin, tau protein, and actin. Misregulation of CDK5 by association with a truncated form of p35 has been implicated in the pathogenesis of Alzheimer's desease.
- **CDK7, Cyclin H**
  The cyclin H-CDK7 (also named CAK, see Section 13.2.2) complex participates in two central regulatory processes, namely the phosphorylation and activation of CDC2 (see Section 13.6) and the phosphorylation of the carboxy-terminal domain (CTD) of RNA polymerase II during transcription initiation (see Section 1.4.3.4).
- **CDK8, Cyclin C**
  One function of cyclin C-CDK8 (review: Sano and Schneider, 2003) appears to be the repression of transcription by RNA polymerase II. Cyclin C-CDK8, which is

found as part of a mediator complex (see Section 1.4.4), phosphorylates cyclin H and thereby inactivates the CTD kinase activity of CDK7.

- **CDK9, Cyclin T**
  Cyclin T-CDK9 (review: Simone and Girodano, 2001) functions as a positive transcription factor during transcription elongation of RNA polymerase II by phosphorylating the CTD. The HIV regulatory protein Tat binds to and specifically activates the cyclin T-CDK9 complex.
- **CDK10**
  This poorly characterized CDK is thought to modulate gene expression by binding to and phosphorylating transcripton factors such as the Ets protein (Kasten and Giordano, 2001).

13.2.2
Structure of CDKs and Regulation by Phosphorylation

The CDKs may exist in inactive and active states. The transition between the two states is controlled by cyclin and/or CKI binding and by phosphorylation/dephosphorylation events (Fig. 13.5). The CDKs possess several phosphorylation sites for protein kinases, and these may have an activating or inactivating effect. Phosphorylation at Thr161 of CDC2 kinase or the equivalent positions Thr160 of CDK2 and Thr172 of CDK4/6 is *activating*. Phosphorylation at Thr14 and Tyr15 is *inhibiting*. A requirement for this regulation to take effect is the association of the CDK with the corresponding cyclin.

Phosphorylation at Thr160 (161, 172): Activation

Phosphorylation of the CycA-CDK2 complex at Thr160 leads to a nearly 300-fold increase in protein kinase activity. Thr160 of CDK2 lies in the activation segment (also known as the T loop, see Section 7.2.1), which blocks the access to the substrate-binding site in the inactive form (see Figs. 13.8 and 13.9). Phosphorylation of the T loop causes – in cooperation with cyclin binding – the T loop to be fixed in a position that affords optimal access of the substrate to the substrate-binding site.

The protein kinase responsible for phosphorylation at Thr160 itself belongs to the family of CDKs and is known as *CAK* (*CDC2 activating kinase*) (review: Kaldis, 1999) composed of the catalytic subunit CDK7 and cyclin H (see Section 13.2.1). CAK demonstrates fairly constant activity during the cell cycle and therefore cannot be the regulating factor for phosphorylation of Thr160 of the CDK. The CAK-catalyzed phosphorylation at Thr160 requires binding of the cyclin to CDK and is regulated by the cyclin concentration.

Phosphorylation at Thr14 and Tyr15: Inactivation

Phosphorylation at Thr14 and Tyr15 leads to inactivation of the CDKs. In the fission yeast, the *Wee1 kinase* is responsible for this phosphorylation, and in mammals there are enzymes homologous to Wee-1 kinase. Phosphorylation at Thr14 and Tyr15 is of particular importance for the regulation of CDK activity in mitosis. The CDC2-cyclin B

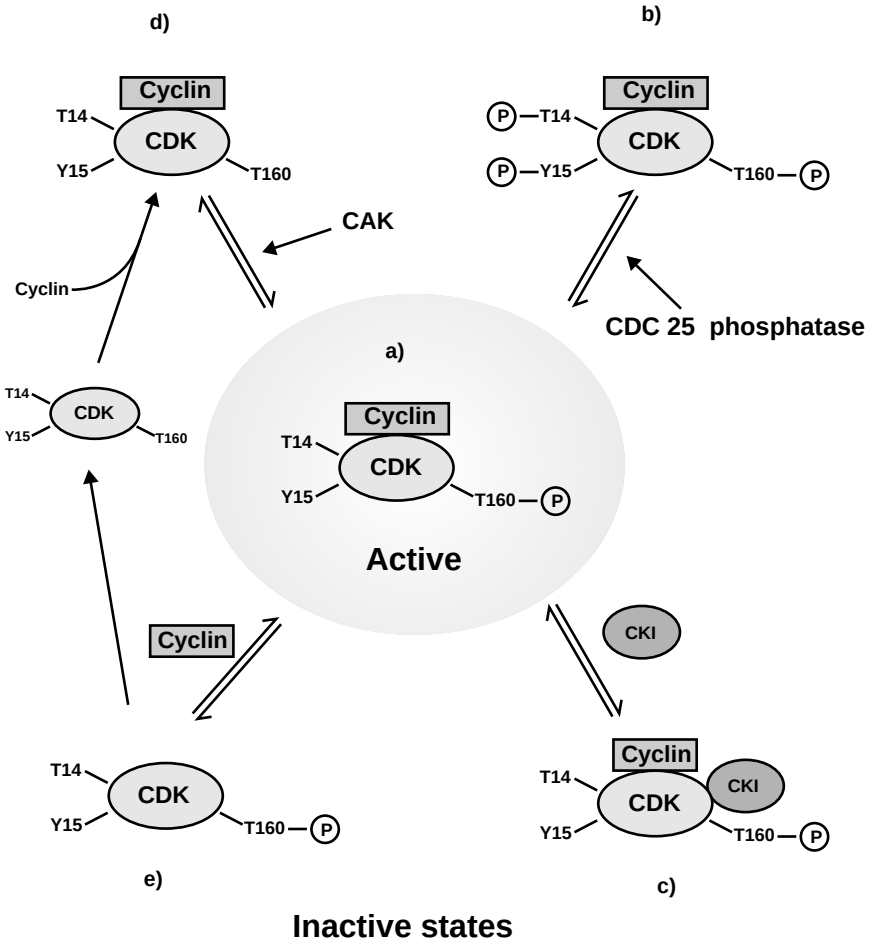


Fig. 13.5 Principles of regulation of cyclin-dependent protein kinases. The figure shows the principles of CDK regulation, using the CDC2 kinase (here simply referred to as CDK) as an example. The active form of CDK (a) is associated with the corresponding cyclin; Thr160 of CDK (or equivalent positions in other CDKs) is phosphorylated, and Thr14 and Tyr15 are unphosphorylated. Inactivation may take place by phosphorylation of Thr14 and Tyr15 (b) or by binding of a CKI (c). Other inactive forms of CDKs are the CDK-cyclin complex, in which Thr160 of the CDK is not phosphorylated (d). In addition, the cyclin-free forms of CDK are inactive (e). CDK: cyclin-dependent protein kinase; CKI: inhibitor of CDK; CAK: CDC2 activating kinase.

complex is maintained in an inactive state until the end of G_2 phase by the phosphorylation of Thr14 and Tyr15. At the G_2/M transition, the inactive state is ended by the action of CDC25 phosphatase, which cleaves off the inhibitory phosphate residues (see below and Section 13.6).

Dephosphorylation at Thr14 and Tyr15: Activation

The inactivating phosphorylation at Thr14 and Tyr15 can be reversed in a regulated manner by CDC25 phosphatase. This enzyme, first described for *S. pombe*, is a protein phosphatase with two-fold specificity that can cleave phosphate residues from phosphoserine and phosphotyrosine residues of CDKs or other proteins. Three different CDC25 phosphatases, CDC25A, B and C have been identified in higher eucaryotes. CDC25A participates in the control of the G_1/S transition and has phosphorylated CDK2 as substrate (see Section 13.7). The CDC25B and C enzymes act at the G_2/M transition by dephosphorylating CDC2 kinase and are thus an essential controlling part of the G_2/M transition (see Section 13.6). Control of CDC25 enzymes is achieved mainly by phosphorylation at specific Ser/Thr residues. Depending on the identity of the Ser/Thr residues, the phosphorylation can have an activating or inactivating effect on the phosphatase activity. Some phosphorylation sites mediate inhibition of the phosphatase by binding of 14-3-3 proteins (see 13.7). The phosphorylations are catalyzed by cyclin-CDK complexes or by other protein kinases that are part of distinct signaling pathways. This type of regulation is also central to the DNA damage and DNA replication checkpoints of the cell cycle (see Section 13.8).

Structure of CDKs

The catalytic center of the CDKs includes a core of ca. 300 amino acids, which adopt the typical protein kinase fold (see Section 7.2.1 and Fig. 13.8). However, comparison of CDK structures without bound cyclin to active structures of, e. g., protein kinase A shows that the catalytic center and the substrate-binding sites of the CDK alone are not organized in a way that allows binding and phosphorylation of substrate. As for other protein kinases, a catalytically important loop, the activation segment or T loop can be identified within the kinase domain of CDKs, which is the target of activating phosphorylations. Another conserved region that is found in the catalytic core of many members of the CDK family comprises a conserved sequence of 16 amino acids known as the *PSTAIRE region*. The PSTAIRE region forms an α-helix (also named a PSTAIRE helix), which is equivalent to the C-helix of other protein kinases (see Section 7.2.1. This region is involved in binding the corresponding cyclins, and helps to determine specificity of cyclin-CDK binding (see Section 13.2.5).

13.2.3 Cyclins

Active CDKs are heterodimers composed of the CDK subunit and an activating subunit, the cyclins. There is specificity in the interaction of cyclins with CDKs, and distinct cyclin-CDK complexes become active during the cell cycle (see Table 13.1). The cyclins were originally defined as proteins that show cyclic concentration variations during the cell cycle (Fig. 13.6) and thereby activate CDKs differentially during the cell cycle. A classifying feature of the cyclins today is the *cyclin box*, a conserved domain of ca. 100 amino acids. Binding to the corresponding CDK takes place via the cyclin box.

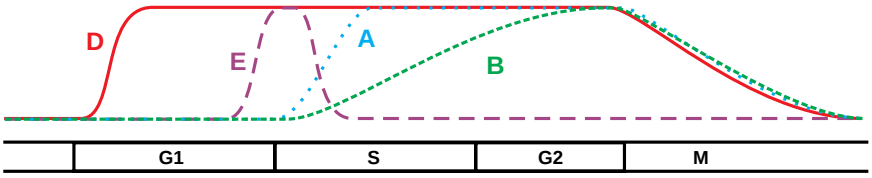


Fig. 13.6 Concentration changes in cyclins during the cell cycle. Cyclins of types D, A, E, and B show characteristic concentration changes in the course of the cell cycle. The figure depicts the approximate course of concentration changes in the cell cycle, and does not indicate the relative cyclin concentrations.

The cyclins in mammals can be roughly divided according to their activity in the different phases of the cell cycle (see also Fig. 13.11):
– The G_1/S cyclins include the D and E type cyclins.
– The *M-phase-specific* cyclins include the B type cyclins.
– Cyclins of type A are active in S, G_2 and M phases.

The equivalent cyclins in yeast have a different classification and are shown in Table 13.1.

The main functions of cyclins are

- **Activation of CDK**
  The role of the cyclins is to convert CDKs into an active state. This process confers specificity to CDK activation, because a specific cyclin preferentially binds and activates only a certain CDK (see Table 13.1). The concentration of the cyclins is an important factor in the control of CDK activity. Various mechanisms exist to control the level of cyclins available for CDK binding (see below).
- **Contribution to substrate specificity of CDKs**
  Binding of protein substrates to CDK-cyclin complexes is not restricted to CDKs. Rather, cyclins contain structural elements that mediate interactions with CKIs and with CDK substrates. Thus, the role of cyclins has to be extended to selection of binding substrates of CDKs.

13.2.4 Regulation of Cyclin Concentration

In the cell cycle, the different cyclins show characteristic concentration changes in which temporally defined maxima in cyclin concentration are observed (see Fig. 13.6). The amount of a distinct cyclin available for CDK activation is strictly controlled by the following mechanisms (Fig. 13.7):

- **Regulation of cyclin expression** (see Section 13.4)
- **Targeted degradation in the ubiquitin pathway** (see Section 13.3.1)
  Most cyclins are the target of ubiquitin-mediated proteolysis, and this degradation is a major mechanism for reducing cyclin concentrations at distinct cell cycle stages.

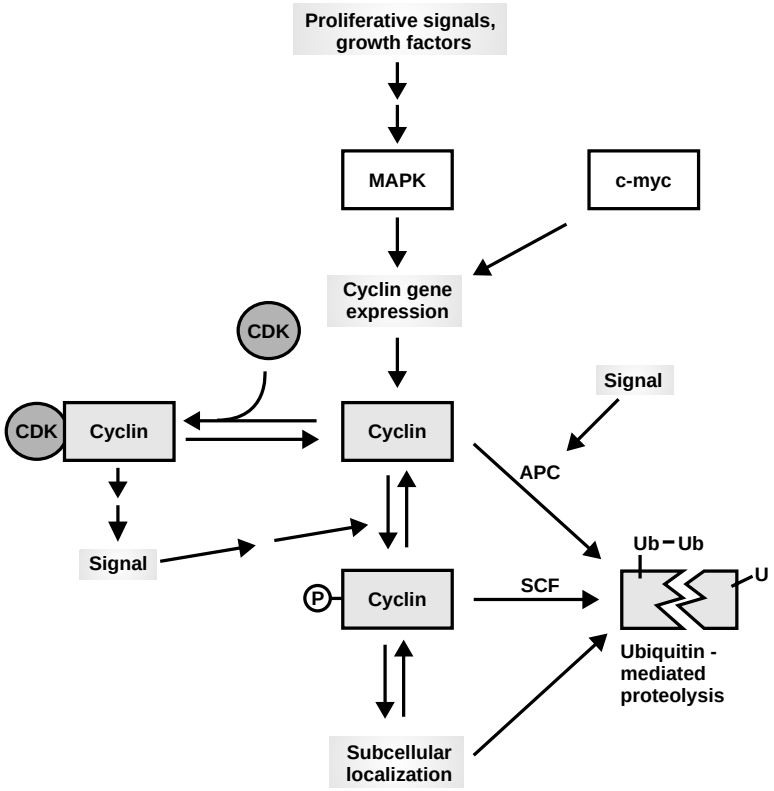


Fig. 13.7 Processes influencing cyclin concentration. Upregulation of cyclin concentration occurrs mainly at the level of gene expression. Growth factors transduce signals via the MAPK pathway or other pathways to transcription factors (c-Jun, c-Myc) that activate e. g. cyclin D transcription. Phosphorylation of cyclins (e. g. cyclin E) can provide a signal for binding of the SCF complex and destruction by the ubiquitin-proteasome pathway. This phosphorylation can be induced by activated cyclin-CDK complexes. Cyclin destruction (e. g. cyclin B) can be also mediated by the anaphase promoting complex. Another mechanism for control of cyclin concentration uses changes in the subcellular distribution.

- **Phosphorylation**
  Signal-directed phosphorylation of cyclins is a tool for targeting cyclins for proteolysis. Furthermore, cyclin phosphorylation influences the subcellular distribution of cyclins.
- **Subcellular distribution**
  Cyclins show a complex pattern of subcellular distribution. Cyclin D1 for instance is localized to the nucleus during G_1 phase and is distributed to the cytoplasm upon the onset of S phase.

13.2.5
Structural Basis for CDK Activation

Binding of the corresponding cyclin and activating phosphorylation are required for full activation of the CDKs. Without the cyclin, the CDKs are inactive; the cyclin-CDK complex possesses a basal protein kinase activity that is considerably increased by phosphorylation at Thr in position 160 (or equivalent position). The structural changes that are the basis of the different states of activity have been worked out for free CDKs, cyclin-CDK complexes, phosphorylated cyclin-CDK complexes, and peptide substrates bound to cyclin-CDK complexes. Most of the structural data could be obtained for the mammalian CDK2 enzyme (reviews: Johnson et al., 1996; Endicott et al., 1999).

Cyclin A-CDK2 Complex

CDK2 has a folding pattern similar to that of the other structurally characterized Ser-specific protein kinases. Comparison of the inactive form of CDK2 (Fig. 13.8) with the cyclin A-bound form and with the active form of protein kinase A shows that there are two main causes of the inactivity of CDK2 without bound cyclin. Firstly, in the inactive form, the binding site for the protein substrate is blocked by a loop known as the T-

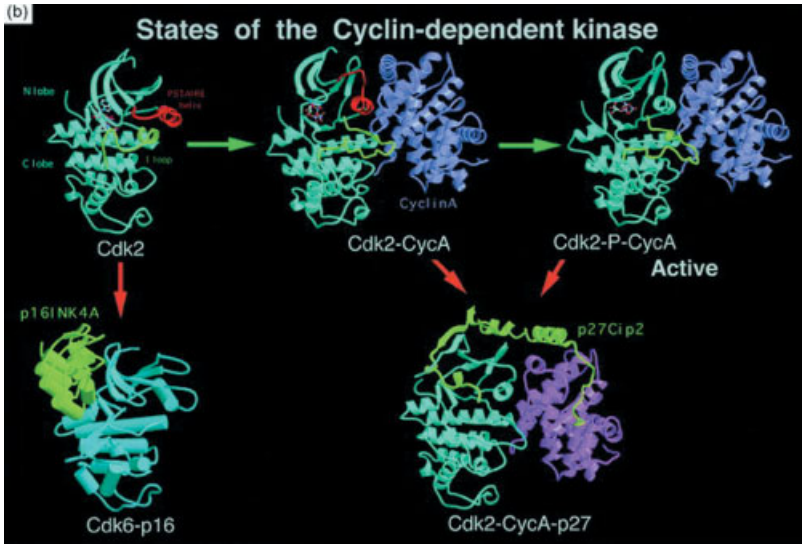


Fig. 13.8 Active and inactive Cyclin-dependent protein kinase structures. The structures shown correspond to non-activated (monomeric CDK2), partially active (CDK2-cyclinA complex), fully activated (Thr160-phosphorylated CDK2-cyclinA complex) inhibited (p27-CDK2-cyclinA complex) and non-activatable (p16-CDK6) complex) states. In the monomeric CDK2 and cyclinA-bound structures, the PSTAIRE helix is highlighted in red and the T loop in yellow. In the phosphorylated CDK2-cyclinA complex, the T-loop is in yellow and the phosphate group is indicated by a yellow sphere. Where present, ATP is shown in ball-and stick-representation. From Pavletich (1999), with permission.

loop, which is equivalent to the activation segment described in Section 7.2.1. The Thr160 important for regulation is found in this T loop; phosphorylation of Thr160 is required for full activation of CDK2. Secondly, the residues of the active center involved in ATP cleavage are oriented in the inactive state, so that cleavage of bound ATP is impossible. Release from the inactive state upon binding of cyclin A leads to a reorganization of the active site, allowing for productive binding of ATP. Furthermore, the T-loop is partially removed from the catalytic cleft and the Thr160 important for regulation is now accessible for phosphorylation by CAK.

Thr160 Phosphate: the Phosphorylated Cyclin A-CDK2 Complex

Phosphorylation of Thr160 by the CAK is accompanied by a ca. 300-fold increase in protein kinase activity. The structure of the phosphorylated cyclin A-CDK2 complex (Fig. 13.9A) differs considerably from the unphosphorylated form at certain critical points. These differences affect the C-terminal lobes of CDK2, the CDK-cyclin A interface, and the T-loop. The Thr160-phosphate serves as the organizing center that contacts different structural elements of the complex and structurally reorganizes them. Thanks to its polyvalent coordination sphere, the Thr160 phosphate couples parts of the T-loop, the catalytic loop with essential Asp127, the PSTAIRE helix , and residues of cyclin A (Fig. 13.9 B). Contacts are formed to Arg50 of the PSTAIRE helix, to Arg150 of the T loop, and to Arg126; the latter lies in the vicinity of the essential Asp127. Furthermore, Van der Waals bonds are formed to an Ile270 of cyclin A. The conformational changes induced by these contacts affect the putative substrate-binding site and the cyclin A-CDK2 interface, in particular.

It is assumed that the activity increase is mainly due to better accessibility of the binding site for the protein substrate. In the unphosphorylated form, the T loop blocks access to the substrate-binding site, whereas in the phosphorylated form this site is exposed.

Substrate Binding by CDKs

The X-ray structures of CDKs with bound peptide substrates have revealed that the substrates bind in an extended form across the catalytic site. The conformational change of the T loop observed upon phosphorylation of Thr160 is critically important to substrate binding.

Substrate binding involves interactions with both the CDK and the cyclin. Whereas the CDK contacts mainly the residues surrounding the target Ser/Thr of the substrate, interactions between distinct sequence elements of the cyclin and the substrate have been shown to contribute to substrate specificity of cyclin-CDK complexes as well. The region of the cyclin responsible for this interaction is also involved in the binding of CDK inhibitors and other regulatory proteins such as the retinoblastoma protein, pRb (see Section 13.4.2).

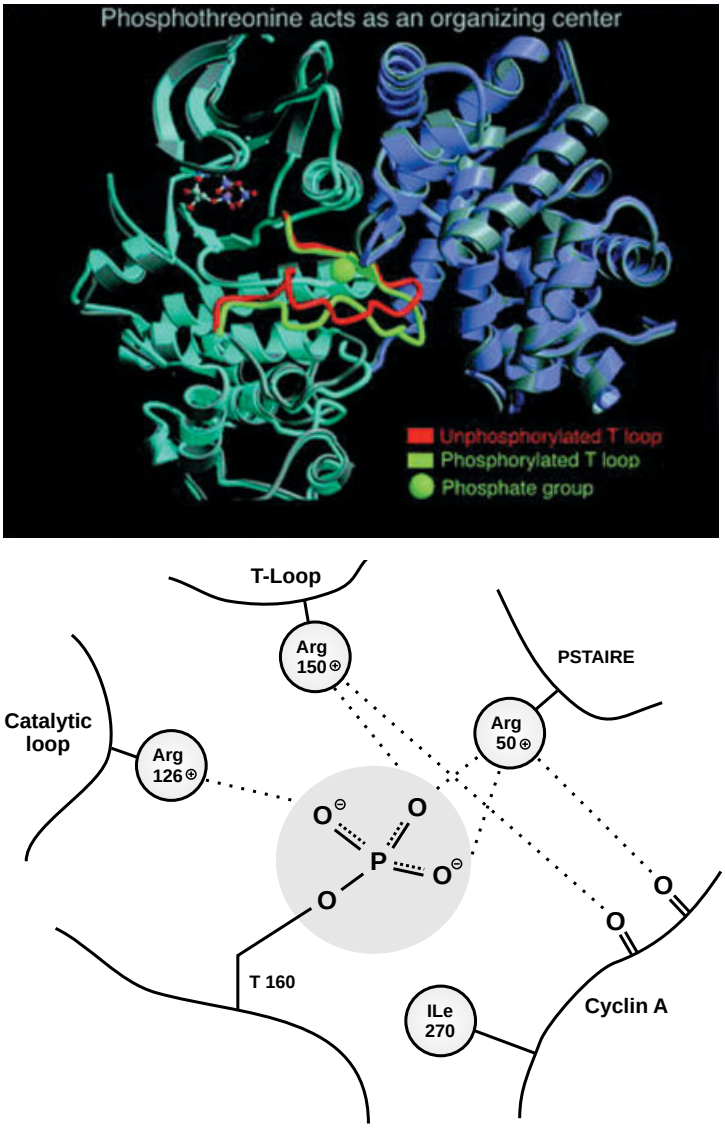


Fig. 13.9 Structure of the Thr160 phosphorylated CDK2-CyclinA complex.
A). Phosphorylation causes a change in the conformation of the T-loop, highlighted in the superposition of the unphosphorylated (gray) and phosphorylated (cyan and magenta) CDK2-Cyclin A complexes. The unphosphorylated T-loop is in red, and the phosphorylated loop in yellow. The phosphate group is indicated by a yellow sphere and ATP is shown in ball-and-stick representation. From Russo et. al. (1996), with permission. .
B) Diagram of the multivalent interactions of Thr-P 160 at the CDK2-Cyclin A interface. Thr-P 160 forms contacts to Arg50 of the PSTAIRE helix, to Arg150 of the T-loop, to Arg126 which is close to the catalytic Asp127 and to Ile270 of Cyclin A. Furthermore, backbone carbonyls of cyclin A help to orient Arg150 and Arg50. From Johnson and Reilly (1996), with permission.

13.2.6 Inhibitors of CDKs: the CKIs

Negative control of CDK activity in the cell cycle is performed by specific inhibitor proteins known as *cyclin-dependent kinase inhibitors, CKIs* (review: Lee and Yang, 2001). These are a heterogenous family of proteins that may associate with a CDK or with a cyclin-CDK complex in a reversible manner, inhibiting CDK activity.

The CDKs are divided into two groups based on sequence homology:

CIP/KIP family	$p21^{CIP1}$ (also known as CIP1, Waf1)
	$p27^{KIP1}$ (KIP1)
	$p57^{KIP2}$
INK4 family	$p15^{INK4b}$
	$p16^{INK4a}$
	$p18^{INK4c}$
	$p19^{INK4d}$

There is specificity in the inhibition of the various cyclin-CDK complexes (Fig. 13.10).:

The members of the CIP/KIP family mainly act on CDK2 complexes and have little effect on the activity of cyclin D-CDK complexes *in vivo* (. The inhibitors of the INK4 family prefentially bind to and inhibit CDK4 and CDK6 complexes.

Mechanism of Inhibition

Crystal structures of CKIs bound to monomeric or heteromeric CDKs have revealed different modes of inhibition for the CIP/KIP and INK4 families. The structure of a ternary complex composed of cyclin A, CDK2 and $p27^{KIP1}$ (Fig. 13.8) shows that the inhibitor interacts with both cyclin A and CDK2 (Russo et al., 1996). Inhibition of kinase activity is explained by alignment of structural elements of $p27^{KIP1}$ in the ATP-binding site of CDK2. This breaks up the glycine-rich phosphate-binding loop. In addition, the ATP-binding site is completely filled by residues of the inhibitor, so that ATP binding is no longer possible. A similar functional principle is likely for the related inhibitor $p21^{CIP1}$.

INK4 inhibitors possess a common structural motif, the ankyrin repeats. The ankyrin repeats have been identified as protein-protein interaction motifs composed of a helical hairpin and two β-motifs. Generally, the INK 4 proteins compete with D-type cyclins for binding to the CDK. The $p16^{INK4a}$ binds to monomeric CDK6 at a site opposite to the cyclin-binding site and allosterically blocks binding of cyclin D. Furthermore, the ATP-binding site of CDK is deformed by the bound inhibitor (Russo et al., 1998).

Regulation and Function of Cell Cycle Inhibitors

The CKIs are important control elements that regulate the G_1/S transition and the transition of cells from a quiescent to a dividing state and vice versa. Primary targets of CKIs are the CDKs, either free or complexed with cyclins, and the main function of CKIs is that of a negative regulator of CDK activity. However, a complex interaction

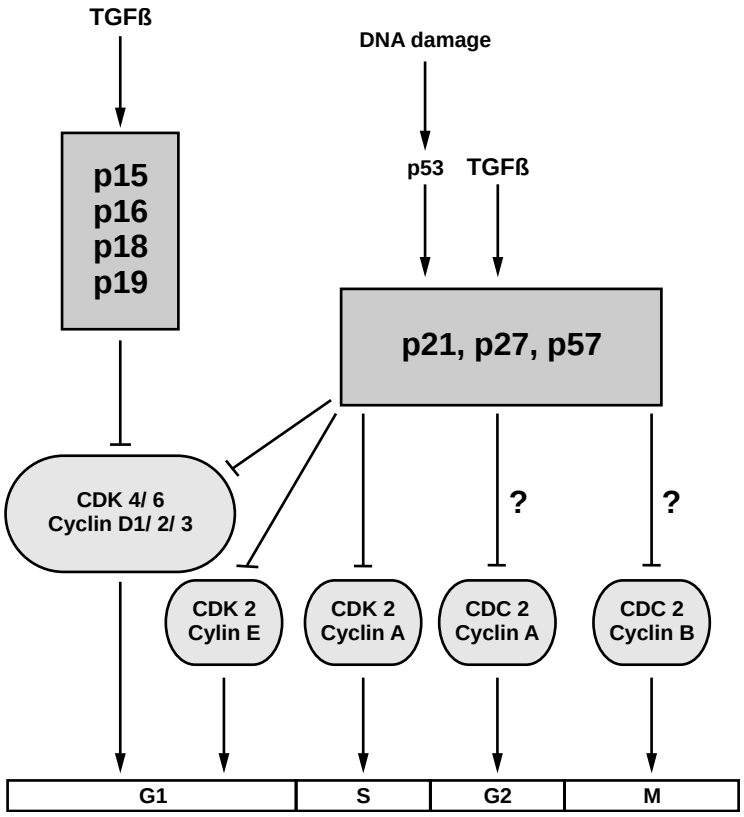


Fig. 13.10 Activation and targets of inhibitors of cyclin-dependent protein kinases (CKIs) in mammals. The Ink inhibitors are activated by TGFβ and specifically inhibit the cyclinD/ CDK4/6 complexes. The inhibitors p21, p27 and p57 are activated e. g. by p53 and TGFβ and can inhibit all types of cyclin/CDK complexes. Only the major ways of CKI regulation are indicated. p53: tumor suppressor protein p53; TGFβ: transforming growth factor β.

network has been found to exist between the CKIs and the CDKs, and we still have incomplete knowledge of the precise functions and regulation of the CKIs.

Regulation of CKIs occurrs by transcriptional, translational, proteolytic and localizational mechanisms, and multiple input signals have been identified that influence these processes in a complex relationship.

The main function and regulation of the CKIs can be summarized as follows:

- **$p27^{KIP1}$**
  This inhibitor appears to play a central role in the decision of a cell to either commit to the cell cycle or to withdraw into the resting state, G_0. The concentration of $p27^{KIP1}$ decreases sharply when cells transit from the quiescent state into the cell cycle, and conversely, levels of $p27^{KIP1}$ increase when cells leave the cell cycle

and enter into a differentiated state. Accumulation of $p27^{KIP1}$ is induced by many antimitogenic signals including cell-cell contacts, TGFβ-signaling and cAMP. The main targets of $p27^{KIP1}$ are the cyclin E-CDK2 and cyclin A-CDK complexes. Surprisingly, $p27^{KIP1}$ is required for the assembly of cyclin D/CDK4 complexes. The role of this property in the physiological context is uncertain.
A major control of the level of $p27^{KIP1}$ is exerted by ubiquitin-mediated proteolysis, which has been shown to be dependent on phosphorylation of $p27^{KIP1}$ by cyclin E-CDK2 complexes (see Section 13.3.1).

- **$p21^{CIP1}$**
  The inhibitor $p21^{CIP1}$ binds mainly to the complexes of CDK2 with cyclin A and E, leading to their inactivation and cell cycle arrest. The outstanding property of $p21^{CIP1}$ is its regulation at the transcriptional level by the tumor suppressor protein p53 (see Section 14.8.3) during the DNA damage response. Increased levels of mRNA for $p21^{CIP1}$ are observed upon treatment of cells with DNA-damaging agents, and this transcriptional regulation is part of the DNA damage checkpoint during G_1/S phase (see Section 13.8). In addition to its ability to associate with G_1 cyclin-CDK complexes, $p21^{CIP1}$ also associates with the replication-accessory protein PCNA (Luo et al., 1995). PCNA (proliferating cell nuclear antigen) is required for nuclear DNA synthesis and functions in clamping DNA polymerase δ to the DNA, thereby increasing the processivity of DNA synthesis. It is believed that this is another mechanism by which $p21^{CIP1}$ can inhibit DNA synthesis and S-phase progression.
- **INK4 proteins**
  The members of the INK4 family bind predominantly CDK4/6 complexes, preventing association of D-type cyclins. Regulation of INK4 abundance is cell-type specific and shows complex patterns. Notably, the level of $p15^{INK4b}$ is subject to induction at the mRNA level by the antimitogenic cytokine TGFβ (see Section 12.1).
  Properties as a tumor suppressor are attributed to the inhibitor $p16^{INK4a}$, since the gene for $p16^{INK4a}$ is mutated in many tumor cell lines.

13.2.7 Substrates of CDKs

Because of the problems in identifying cellular substrates of protein kinases, as described in Chapter 7, it has been a difficult and lengthy process to determine the functionally relevant substrates. Figure 13.11 gives an overview of the cell-cycle-specific activation of CDKs and some important substrates. Comparatively sparse information is available on the G_1- and S-phase substrates of the CDKs. In contrast, many proteins have been described that undergo specific phosphorylation in G_2/M phase. The sequence (K/R)-S/T-P-X-K (X: any amino acid) has been identified as a consensus sequence for phosphorylation by CDKs.

Substrates in G_1/S Phase

The most important CDK substrate in G_1/S phase are the tumor suppressor protein pRb and the pRb-related proteins p130 and p107, which are phosphorylated by the

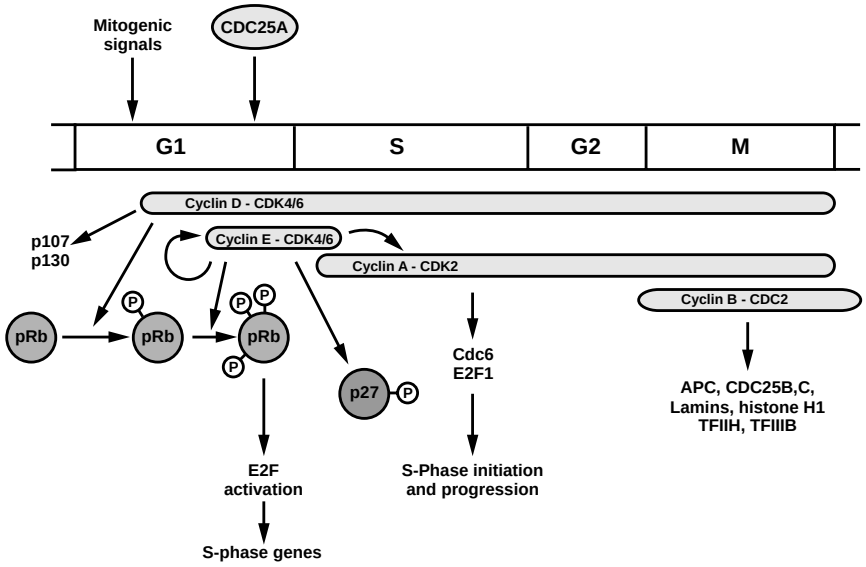


Fig. 13.11 Substrates and phase-specific activation of CDKs in the cell cycle. An overview is shown of the phase-specific activation of the most important CDK-cyclin complexes and of selected substrates. The arrows indicate activation and phosphorylation. p107, p130: pRb related proteins; NPAT (nuclear protein mapped to the ATM locus); RSK: ribosomal S6 kinase members; APC: anaphase promoting complex; for MCM proteins see 13.5; for Cdc20 see 13.3.2; for Skp2 see 13.3.1.

cyclin D1-CDK4/6-complex and by the cyclin E-CDK2 complex. The protein pRb and its relatives are critical components for crossing the restriction point and entry into S phase. Other targets of the cyclin E-CDK2 complex comprise the $p27^{KIP1}$ inhibitor and MCM proteins (see Section 13.5), which are involved in the regulation of DNA replication.

The cyclin A-CDK2 complex has been shown to phosphorylate, among others, the transcription factor E2F1 and components of the DNA replication complex.

Substrates in G_2/M Phase

In M phase, new phosphorylation of many proteins is observed that starts, in particular, from the CDC2-cyclin B complex. The phosphorylation mostly affects proteins involved in the reorganization of the cytoskeleton, the nuclear membrane, and the formation of the spindle apparatus. As a consequence of phosphorylation events, inhibition of vesicular transport and general inhibition of transcription occur.

Examples of proteins that are specifically phosphorylated during the M phase of the cell cycle are the lamins. Hyperphosphorylation of the lamins leads to disintegration of the nuclear lamina. A myosin-associated protein named MAP4 is also specifically phosphorylated during mitosis. Other M-phase-specific phosphorylations occur at transcription factor TFIIIB, leading to inhibition of transcription by RNA polymerase III. Phosphorylation of TAF proteins (see Section 1.4.3.3) is also involved in general inhibition of transcription.

13.2.8
Multiple Regulation of CDKs

Activity of the regulatory components of the cell cycle varies extremely during the course of the cycle, and it is directed by external signals and internal control mechanisms. The CDKs are the central tool for control of the cell cycle. They receive a multitude of signals and transmit signals to downstream substrates. These substrates trigger the large number of different activities that constitute the distinct functional, regulatory and morphological properties of the various phases of the cycle. The activity of the CDKs is regulated by various positively and negatively acting signals that are registered in the cell cycle and can bring about a halt in the cell cycle at various points.

The following have a *positive* effect on activity of CDKs and promote progression in the cell cycle:

- increase in cyclin concentration by activation of transcription or inhibition of proteolytic degradation
- phosphorylation of CDKs at Thr160 or equivalent positions by CAKs
- dephosphorylation of CDKs at Thr14/Tyr15 by Cdc25 phosphatases
- redistribution of CKIs between different CDK complexes
- increase in concentration of CDK4 and CDK6
- decrease in concentration of CKIs at the transcription level or by proteolysis.

The following have a *negative* effect and can lead to a stop in the cell cycle:

- decrease in concentration of cyclins by reduced transcription or by activation of proteolysis
- phosphorylation at Thr14/Tyr15
- increase in concentration of CKIs.

These regulation mechanisms cannot be considered in isolation. Rather, it must be assumed that the individual mechanisms cooperate, that they demonstrate mutual regulation, and that feedback mechanisms are built in. All control elements can be activated, in principle, by external signals, resulting in a complex network of cell cycle control with many entry and exit points. The following sections are thus incomplete and only describe the elements that have been well proven experimentally.

13.3
Regulation of the Cell Cycle by Proteolysis

The ordered course of the cell cycle is ensured by two processes in particular:

- regulated and temporally coordinated activity changes in CDKs
- targeted ubiquitin-dependent proteolysis of CDK regulators.

Both processes, the protein kinase regulatory network of CDKs and targeted proteolysis, are linked to one another and work in mutual dependence. For a review on the control of the cell cycle by targeted proteolysis see Tyers and Jorgensen, 2000; Yew, 2001).

Selection of a protein for Ub-dependent proteolysis occurs particularly via the E3 enzymes of the ubiquitin pathway (see Section 2.6.1) that catalyzes ligation of the target protein with ubiquitin. The specificity of Ub-protein ligation is determined by the nature of the E3 enzymes, which constitute a large family of functionally related but structurally distinct proteins of heterogeneous subunit composition.

Use of specific proteolysis as a tool for control of the cell cycle has various advantages:

- Proteolysis allows simultaneous and complete inactivation of all functions of a multifunctional cell cycle protein such as the cyclins.
- Proteolysis enables subunit-selective reorganization of hetero-oligomeric protein complexes. An example is the targeted degradation of a CDK inhibitor.
- The total substrate pool of regulatory enzymes of the cell cycle may be inactivated by proteolysis.
- The regulatory system of a cell cycle section can be reset to the ground state by proteolysis.

Two types of E2/E3 complexes are of particular importance for cell cycle control (see Fig. 13.12). One, the SCF complex, is of outstanding importance for the G_1/S transition. The other, the anaphase-promoting complex (APC) or cyclosome, is especially important for the course and control of mitosis. Common to both complexes is the variable collaboration with different proteins to mediate the ubiquitinylation of different substrates.

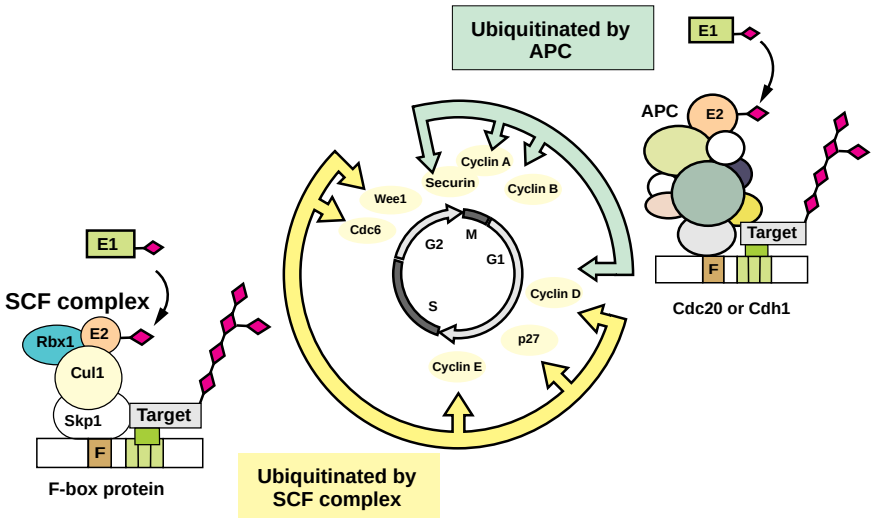


Fig. 13.12 Roles of two distinct ubiquitin ligases in regulation of the cell cycle. The ordered progression of the cell cycle is regulated by two ubiquitin ligases: the anaphase promoting complex/ cyclosome (APCC) and the SCF complex. The APC is active from late G2 to mid-G1 phase and catalyzes the ubiquitination of mitotic cyclins and securins which are anaohase inhibitors. In contrast, the SCF complex mediates ubiquitination of G1 cyclins and CKIs. After Nakayama, 2001.

13.3.1 Targeted Proteolysis by the SCF Complex

Progress of the cell cycle through G_1 phase, crossing of the restriction point R, and entry into S phase are associated with the ubiquitin-dependent proteolysis of important regulatory cell cycle proteins. Among the proteins that are targeted for ubiquitin-dependent proteolysis are both positive and negative regulators of G_1 and S phases. Although ubiquitin-dependent proteolysis has been demonstrated for a variety of cell cycle regulators in G_1 and S phase, it is not yet clear what degradation events are necessary for, e. g., progression into S phase in vertebrates. It also remains to be established how the degradation of proteins may function to order the complex events during G_1 and S phase. Nevertheless, ubiquitin-dependent proteolysis has been recognized as an important part of progression through G_1 and S phase in vertebrates.

The regulators that are targeted for degradation include proteins whose degradation may be a requirement for crossing the restriction point and entry into S phase as well as those regulators whose degradation may be critical for the maintenance of proper cellular homeostasis during the cell cycle (Yew, 2001). Examples of the first category include the inhibitors p27KIP1, p57KIP2, p21 CIP1 and cyclin A, while the second category includes cyclin E, cyclin D, and the Cdc6 protein (see Section 13.5).

Ubiquitinylation during G_1 and S phase is mediated by E2/E3 enzymes organized in multiprotein complexes called SCF complexes (Skp1, cullin, F-box protein). In mammals, this complex is composed of the protein Skp1 which has a scaffolding function, a family of proteins named cullin, and a small protein Rbx1. The Rbx1 protein (also named Roc1) contains a RING finger motif (see Section 2.6.1) which mediates binding to the E2 enzyme, usually Cdc34. The core of the SCF complex associates with a family of proteins named F-box proteins that function as substrate-specific adaptor subunits and confer substrate specificity by recruiting a particular target to the core ubiquitination machinery. The F-box proteins have a sequence element named F-box that mediates binding to Skp1, and they capture phosphorylated substrates by means of carboxy-terminal protein-protein interaction regions such as WD repeats or Leucine-rich motifs (review: Peters, 1998). Figure 2.15 gives an overview of one mammalian SCF complex containing the F-box protein Skp2 with the assumed functions and contacts during ubiquitination.

Nearly 50 different F-box proteins have been identified in mammals, and consequently multiple forms of SCF complexes exist because of this variable association (see Fig. 13.13).

Some of the ubiquitin-protein ligases involved act on their specific substrates only after phosphorylation of the substrate protein. Examples include the inhibitor $p27^{KIP1}$, which requires phosphorylation by cyclin E/CDK2 on Thr187 for binding to the Skp2-SCF complex. Phosphorylation of $p27^{KIP1}$apparently has a twofold function during Ub-mediated degradation (Tomoda et al., 1999). To be broken down, $p27^{KIP1}$ must be transported out of the nucleus, which requires its phosphorylation. Furthermore, phosphorylation is needed for recognition by the ubiquitin-conjugating system. Export from the nucleus and proteasome-mediated degradation are both controlled by phosphorylation in this case. Thus, the regulator for this degradation pathway

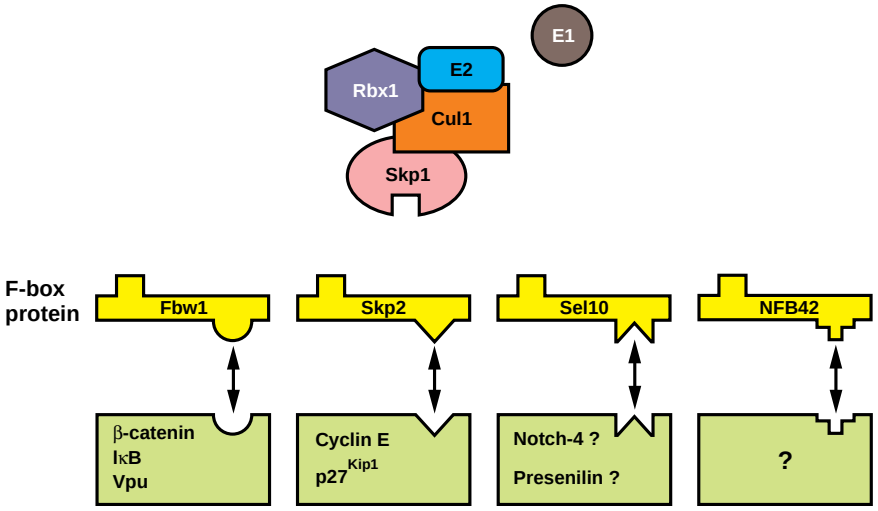


Fig. 13.13 Role of the SCF complex as a versatile E3 ubiquitin ligase. The SCF complex is composed of the proteins Skp1, Cul1, The RING-finger protein Rbx1 and an F-box protein. Each F-box protein contains an F-box domain, which is required for binding to Skp1, as well as a specific substrate recognition motif. Ubiquitin is transferred from E2 to the substrate with the help of the SCF complex. Whereas Skp1, Cul1 and Rbx1 are invariable components of the SCFcomplex, F-box proteins are highly variable and interchangeable. The Fbw1 subunit directs ubiquitination towards β-catenin (see Section 14.9), the inhibitor IκB (see Section 2.6.4) and protein Vpu from the human immunodeficiency virus Typ 1. The Skp2 subunit recognizes cyclinE and the CDK inhibitor $p27^{KIP1}$ and Sel10 possibly interacts with presenilin and notch (see Section 12.2).

is the phosphorylation of the substrate by a regulatory protein kinase, and the Ub-ligase system may be constitutively active.

Many substrate proteins are selected for Ub ligation based on a C-terminal target sequence. These sequences which, because of the occurrence of common amino acids, are known as PEST sequences, are often targets for phosphorylation.

13.3.2 Proteolysis during Mitosis: the Anaphase-promoting Complex/Cyclosome

The anaphase-promoting complex (APC, also named cyclosome) is another type of ubiquitin-ligase that mediates the proteolysis of important regulators of cell cycle progression, with a major effect during mitosis (reviews: Page and Hieter, 1999; Harper and Solomon, 2002). The activity of the APC is tightly regulated to control cell cycle progression, being high from late mitosis until late in the G_1 phase, but low in S, G_1, and early mitosis in mammalian cells. A multitude of APC substrates have been identified including cyclin A, cyclin B, mitotic protein kinases, inhibitors of anaphase, spindle-associated proteins, and inhibitors of DNA replication. One hallmark of APC is its complicated structure: it is composed of at least 11 different subunits that assemble into a structure reminiscent of the proteasome structure. In APC,

an outer protein wall surrounds an inner cavity where ubiqutinylation is thought to take place (Gieffers et al., 2001).

It is largely unknown which subunits of APC are involved in recruiting the E2 enzyme and in selecting the substrate. At least part of the substrate recognition appears to be due to the reversible association of the APC with activating subunits, of which two types have been identified: the Cdc20 protein and the Hct1/Cdh1 protein. Both Cdc20 and Hct1 have been shown to associate with specific substrates. Accordingly, different types of APC complexes exist with distinct substrate preferences and with different functions in the cell cycle. It is believed that Cdc20-APC is required both for entry into early anaphase and exit from mitosis. On the other hand, Hct1-APC is assumed to maintain APC activity from the end of mitosis until the end of G_1.

Most substrates for APC-mediated ubiquitinylation carry a particular sequence, the destruction box. For B-type cyclins, the consensus sequence of the destruction box is R-ALGVN/D/EI-N. Deletion of the destruction box in cyclin B causes its stabilization.

Important APC substrates are the mitotic cyclins and the anaphase inhibitors. Proteolysis of mitotic cyclins is activated at the metaphase/anaphase transition and is only switched off at the start of S phase. The mitotic cyclin A is degraded before cyclin B. Lasting activity of the Hct1-APC during G_1 phase is thought to be responsible for the lack of detection of mitotic cyclins in G_1 phase. Renewed accumulation of mitotic cyclins is only possible again when APC activity is switched off at the start of S phase.

Another class of substrates comprises the anaphase inhibitors, e. g., the protein securin. Securin is an inhibitor of a protease named separase, which cleaves proteins responsible for sister chromatid cohesion in metaphase. Destruction of these inhibitors is necessary for triggering of sister chromatid separation and progression into anaphase. Because of its central function, the APC is part of several cell cycle checkpoints, e. g., a DNA damage checkpoint and the mitotic spindle checkpoint.

Regulation of APC activity is complex and is performed by various mechanisms. It involves the variable association of the activating subunits, which may be inactivated by binding to specific inhibitors. Another important element of APC regulation is phosphorylation of specific subunits or APC regulators. However, because of the complexity of the APC system it is largely unknown how phosphorylation signals feed into the ordered APC function during the cell cycle.

13.4 The G_1/S Phase Transition

G_1 phase has a special regulatory function in the cell cycle; here, the decision is made to enter S phase and thus a new round of cell division, or to enter a resting, quiescent state.

When mitosis has been completed, the cell requires signals in the form of growth factors to direct towards a new round of division. The signals become effective in the early part of G_1 phase. In this time window, the cell is programmed to begin a new cell cycle or to enter G_0 phase. After a particular point, the restriction point R, no further

signals are needed to continue the cell cycle. The cell cycle apparatus is self-contained from this point onwards. S, G_2, and M phase occur without external control. The cell cycle may still be halted after crossing the restriction point, however, if the cell detects, via internal control mechanisms or checkpoints, that defects have occurred in the correct course of the phases. The restriction point thus divides G_1 phase into a growth factor-dependent early part and a growth factor-independent late part. Whether cells cross R depends on the intensity of mitogenic signaling. If growth factor signaling is insufficient, cells exit the cell cycle and enter into G_0, the quiescent state.

13.4.1 Function of the D-type Cyclins

Progress of the cycle in G_1 phase is controlled in particular by the *cyclins of types D and E* and by *inhibitors of the CDKs*. The two types of cyclins perform different functions during G_1. Whereas the cylin D-CDK2/4 complexes prepare the cell for crossing R, the decision for passage through R and entry into S phase is critically dependent on the activity of cyclin E-CDK2 complexes (Fig. 13.14).

Cyclin D

Cyclins of type D are of particular importance in the early part of G_1 phase. Of the three cyclins of type D (D1, D2, D3), two (D2 and D3) do not occur in all cell types, whilst cyclin D1 has a central function in the regulation of G_1 phase in all cell types. *Binding partners* of the D type cyclins include CDK4 and CDK6 in particular, and CDK4 activation is considered to have a key role. For full activation of CDK4, activating phosphorylation at Thr172 of CDK4 is also necessary, in addition to binding of cyclin D1. This step is catalyzed by CAK.

The D-type cyclins complexed to CDK4/6 perform a dual task in G_1 phase regulation. One task is to integrate external signals into the cell cycle. Mitogenic signals, such as growth factors, activate the transcription of the gene for cyclin D1 and thus increase the amount of cyclin D-CDK4/6 complexes. The activating signals can become active from the start of G_1 onward, and a constitutive activity of Cyclin D-CDK4/6 complexes is then observed during the whole of G_1 and during S phase. The increase in cyclin D-CDK4/6 complexes is postulated also to sequester the $p27^{KIP1}$ inhibitor bound to cyclin E-CDK2 complexes away from CDK2. This is thought to establish initial levels of active cyclin E-CDK2, which is necessary for crossing the restriction point.

Another essential function of cyclin D-CDK4/6 complexes is to activate metabolism and to promote cell growth. Cells must exceed a critical size in order to be able to pass through the restriction point R. The best characterized substrate of cyclin D-CDK4/6 complexes is the pRb protein, which is transformed from an unphosphorylated state to a hypophosphorylated state by cyclin D-CDK4/6 complexes.

Cyclin E

Cyclin E binds and activates CDK2. It has now been recognized to be a key regulator of G_1/S transition, functioning in a nonoverlapping way with cyclin D complexes. The

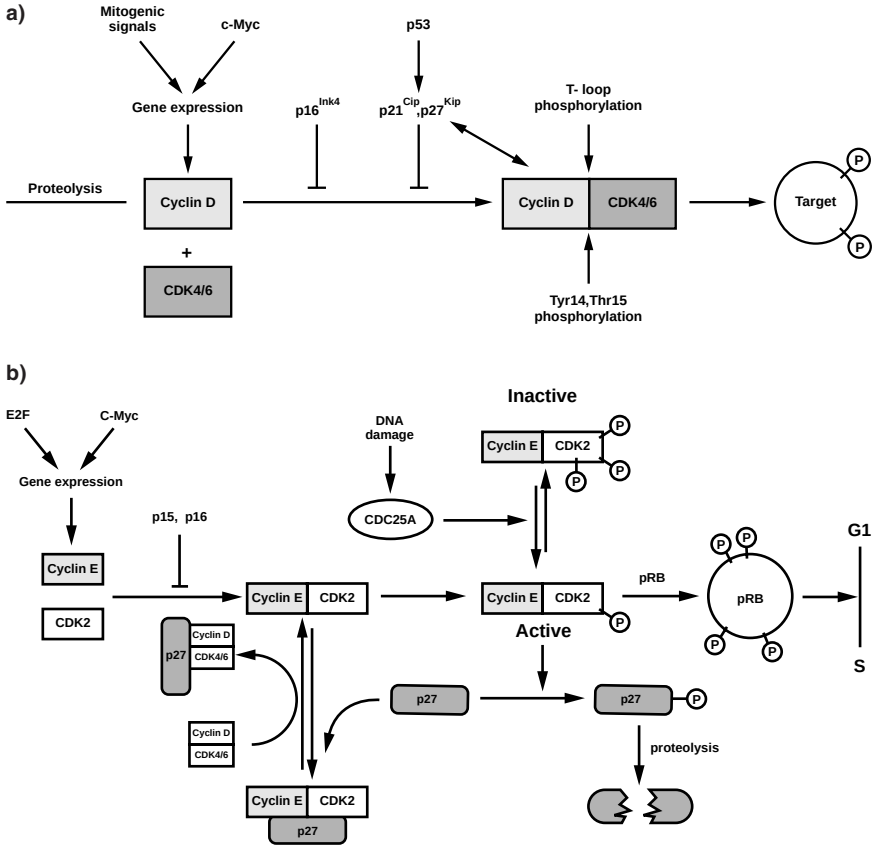


Fig. 13.14 A) Regulation of cyclin D. On the example of cyclin D1, the figure illustrates the activating and inactivating influences on D-type cyclins. For explanation see text.
B) Multiple regulation of cyclin E. The concentration of active cyclin E-CDK2 complexes is determined by expression of the cyclin E gene, by binding of the inhibitors p15, p16 and p27 and by its phosphorylation status. The inhibitor p27 can be removed from p27-cyclin E-CDK2 complexes by sequestering it to cyclin D-CDK4/6 complexes. The amount of active cyclin E-CDK2 is also controlled by the action of CDc25A phosphatase which cleaves off inhibitory phosphate residues from CDK2. Important substrates of active cyclin E-CDK2 complexes are the pRb protein and the inhibitor p27. Phosphorylation of the latter induces its ubiquitinylation and proteasome degradation. Only the major regulatory influences on cyclin E are shown.

cyclin E-CDK2 complexes are activated concomitant with transition across the restriction point, and their activity shows a maximal value at the start of S phase. Afterwards, the cyclin E-CDK2 activity falls off sharply within S phase.

The activity of cyclin E-CDK2 complexes is mainly directed toward two substrates, the pRb protein and the inhibitor $p27^{KIP1}$. Phosphorylation of pRb inactivates pRb and releases the transcription factor E2F from a repressed state, allowing transcription of key proteins for G_1/S transition. The gene for cyclin E is also induced by transcription

factor E2F, which explains the increase in cyclin E at the G_1/S transition. Another important substrate of cyclin E-CDK2 complexes is the inhibitor $p27^{KIP1}$. Phosphorylation of this inhibitor on Thr187 induces its ubiquitination and targets it for degradation in the proteasome.

Negative Regulation of the G_1/S Transition

In addition to mitogenic signals, antimitogenic signals are also processed during G_1 phase. These lead to an increase in the level of CDK inhibitors or they influence the phosphorylation state of the CDKs via the phosphatase Cdc25. Antimitogenic signals can lead to a halt during G_1 phase and bring the cell into a resting state. An antimitogenic signal originates, for example, from TGFβ, from cAMP, from certain cell-cell contacts, and from DNA damage.

Negative regulation of the cell cycle in G_1 phase is performed by the inhibitors of the INK4 family, which preferentially bind and inactivate monomeric CDK4/6 complexes, preventing cyclin D activation. Accumulation of $p16^{INK4a}$, for example, sequesters CDK4/6 complexes, preventing progress in G_1 phase. The inhibitors $p21^{CIP1}$, $p27^{KIP1}$, and $p57^{KIP2}$ are directed mainly against heterodimeric CDK2 complexes and can thereby inactivate the cyclin E/cyclin A-CDK2 complexes, preventing crossing of the restriction point and entry into S phase.

The balance between activated G_1-cyclin-CDK complexes and the various inhibitors controls progress through G_1 phase. The concentration of the inhibitors is regulated in a complex pattern by external cues. Examples of external influences on CKI proteins include the induction of $p21^{CIP1}$ by p53, the stimulation of the degradation of $p27^{KIP1}$ by growth factors, and the induction of $p15^{INK4b}$ by TGFβ. Further reaction and coordination of the positive and negative signals are mediated by a regulation system in which the product of the retinoblastoma gene, the pRb protein, plays a central role (reviews: Planas-Silva and Weinberg, 1997; Adams and Kaelin, 1998).

13.4.2 Function of pRb in the Cell Cycle

The pRb protein is a nuclear phosphoprotein of 105 kDa and belongs to a class of proteins called the pocket proteins. Two relatives of pRb are known, p130 and p107, which share many of its biological properties. The domain structure of pRb is shown in Fig. 13.15 (reviews: Weinberg, 1995; Stiegler and Giordano, 2001). The pRb protein can be divided into at least three functional regions: an N-terminal region, a central pocket, and a C-terminal region. The N-terminal region appears to be required for oligomerization, the pocket region contains binding sites for the transcription factor E2F, for the viral oncoproteins TAg, E1A ,and E7 (see Chapter 14), and for a large number of other cellular proteins. A nonspecific DNA-binding domain is found on the C-terminal region. Numerous Ser/Thr phosphorylation sites have been identified on pRb, and the different phosphorylation events appear to regulate distinct pRb functions.

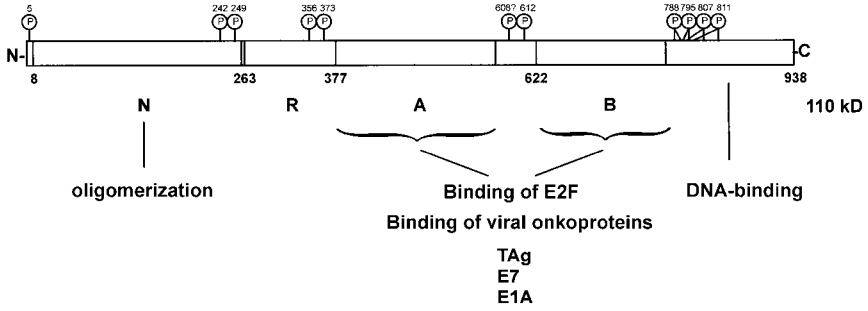


Fig. 13.15 Domain structure of the retinoblastoma protein pRb. The phosphorylation sites (P) of pRb and the localization of the sequence sections necessary for interaction with viral oncoproteins and with the transcription factor E2F are shown. In addition, an oligomerization domain and a DNA binding domain can be identified.

The pRb protein has the characteristics of a *tumor suppressor protein*. Loss of its function is associated with deregulation of cell division and favors tumor formation (see Chapter 14).

A central function of pRb in control of the cell cycle is illustrated by the following:

- At the end of mitosis until the restriction point R, pRb exists in an underphosphorylated form. In the underphosphorylated form, pRb inhibits cell proliferation in that it blocks the activity of transcription factors that control the expression of S phase genes.
- During or after crossing the restriction point, pRb exists in a hyperphosphorylated form, and it remains in this form until the end of mitosis. In the hyperphosphorylated form, pRb has a growth-promoting function.

It is now commonly accepted that mitogenic signals induce – via cyclin-CDK complexes – the phosphorylation of pRB and thereby control the passage through the restriction point and the entry into S phase. The crucial control element of pRb function is its phosphorylation status, which can be considered as a switch. At the start of G_1 phase, pRb exists in an underphosphorylated form. In this form, it functions as a brake on the progress of the cell cycle. The brake is lifted when pRb is phosphorylated by the central components of the cell cycle apparatus. The protein complexes involved in the phosphorylation of pRb are cyclinD-CDK4/6, cyclin E-CDK2, and cyclin A-CDK2/.

Most functions of pRB appear to be linked to its association with a family of transcription factors, named E2F. The protein pRb binds to proteins of the E2F family and thus controls their transcription-activating function.

The E2F Transcription Factors

The E2F proteins are DNA-binding proteins, which, in addition to the DNA-binding domain, also have a binding site for pRb protein and – except E2F4 and E2F5 – a transactivating domain (review: Trimarchi and Lees, 2001). In some cell types, the E2F

proteins exist as heterodimers in complex with another DNA-binding protein known as DP. Eight human genes have been identified as components of E2F transcriptional activity. These genes have been divided into two distinct groups: the E2Fs (E2F1-E2F6) and the DPs (DP1 and DP2). The DNA-binding element assigned to the E2Fs is found in promoter regions of genes that perform an essential function in S phase.

For example, E2F binding sites are found in genes for

- thymidine kinase
- dihydrofolate reductase
- DNA polymerase α
- cyclin A, cyclin E
- transcription factor c-myc
- E2F-1
- PRb
- pro-apoptotic protein Apaf1 (see Section 15.5)
- tumor suppressor ARF (see Section 14.7).

Transcription factor E2F thus controls the expression of proteins that are required for further progress of the cell cycle. Overall, the transcription-controlling activity of E2F can be assigned a central function in the progress of G_1 phase and also S phase.

The relatives of pRb, p107 and p130 bind to E2F members different from that bound by pRb, and these associations occur at distinct stages of the cell cycle. Whereas pRb is associated with E2F both in quiescent and actively dividing cells, p130 binds to E2F predominantly in cells that have entered the G_0 state.

Model of pRb Function

Whereas the overall effect of pRb phosphorylation was recognized quite early, it is now becoming increasingly clear that progressive phosphorylation of pRb on distinct sites controls its growth- and proliferation-inhibiting function in a complicated way. Many aspects of this control are still incompletely understood. Overall, pRb is phosphorylated by activated CDKs at many (>10) Ser/Thr residues. The different CDKs preferentially phosphorylate pRb at distinct sites. There is evidence that the various phosphorylation events have different effects on pRb function.

The present picture of pRb control may be summarized as follows (Fig.13.16):

- In late M and in early G1 phase, pRb exists in an underphosphorylated or even unphosphorylated state. In this state, pRb is associated with E2F-DP members and is bound to promotors of E2F-responsive genes, and these genes are thereby repressed. Repression of E2F activity by pRb appears to be mediated through two mechanisms. A direct repression can be mediated by binding of pRb to the transactivation domain of E2F-DP, preventing its interaction with the core transcription machinery. In another reaction, this active repression may be enforced through the recruitment of protein complexes that facilitate nucleosome packaging. Examples of these proteins include histone deacetylases and histone methyltransferases. Overall, the cooperation of both effects is believed to result in an inactive, repressed state at E2F-responsive genes.

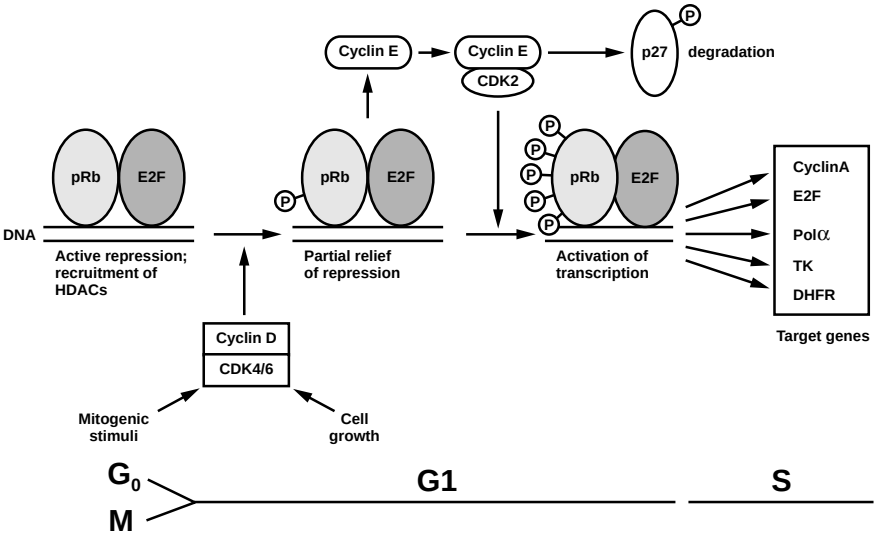


Fig. 13.16 Function of pRb and G1 cyclins upon G_0/G_1 transition and during G_1 progression. During G_0 and in the absence of mitogenic stimuli, pRb is in a hypophosphorylated state and actively represses genes by recruitment of histone deacetylases (HDACs). Mitogenic stimuli induce accumulation of active cyclinD/CDK4/6 complexes which promote cell growth and mediate partial phosphorylation of pRb. Repression of E2F target genes which include the gene for cyclin E, is now partially relieved and cyclin E/CDK2 complexes accumulate completing phosphorylation of pRb. Hyperphosphorylated pRb dissociateas from E2F allowing full transcription of the E2F target genes and progression into S-pahse. The activated cyclin E/CDK2 complex also phosphorylates the inhibitor $p27^{KIP1}$ inducing its ubiquitination and proteasome degradation.

- Progressive phosphorylation of pRb by activated CDKs loosens its grip on the E2Fs and allows transcription of E2F-responsive genes. Under the influence of mitogenic signals, cyclin D-CDk4/6 complexes become activated, and phosphorylation of pRb is initiated, with progressive release of the repressed state. Since the gene of cyclin E is one of the E2F-regulated genes, levels of cyclin E begin to rise, and finally the cyclin E-CDK2 complex is mainly responsible for converting pRb into the hyperphosphorylated form. E2F is now fully released from the inhibitory action of pRB, and full transcription of the E2F-responsive genes is possible. In this state, E2F recruits histone acetylase activities and maintains transcription of its target genes. Distinct roles can be assigned to the various E2F proteins in this process. Whereas E2F1–3 proteins function as genuine transcriptional activators, the E2F4 and E2F5 proteins have only a repressive function. They do not have a transactivation domain and therefore cannot actively promote transcription.

In addition to the E2F proteins, a number of other signaling proteins have been reported to interact with pRb, and it is assumed that pRb is also involved in processes other than the G_1/S control. However, these other potential functions of pRb are not well defined. The *MDM2 protein* has been identified as a further control element that

can influence pRb-E2F function. The MDM2 protein (review: Michael and Oren, 2002) was discovered as an oncoprotein activated by overexpression. It binds to the p53 protein (see Chapter 14) and to pRb protein. By binding, the growth-controlling function of both proteins is stopped. MDM2 has been shown to be a RING-finger (see Section 2.6.1) ubiquitin-ligase E3 enzyme that induces degradation of p53 via the ubiquitin-proteasome pathway. Furthermore, the MDM2 protein also binds to E2F and stimulates its transcription-activating function. Overall, the MDM2 protein therefore has a growth-promoting function. The precise function of the pRb-MDM2 interaction is, however, uncertain.

pRb as Integrator of Positive and Negative Signals

Within the pRb function, there is a meeting of *positive* signals, i. e., those that promote progress of the cell cycle, and *negative* signals that halt the cell cycle.

Positive Signals

The extent of phosphorylation, and thus the activation of pRb, is closely associated with the concentration of the D- and E-type cyclins and the activity of the corresponding CDK complexes. External signals can promote progress in the cell cycle by increasing the activity of these complexes, promoting pRb phosphorylation and release of E2F from the inhibited state, allowing transcription of S phase genes.

A further influence on pRb function may be performed by viral oncoproteins. These bind to the pocket region of underphosphorylated pRb and compete with E2F for pRb binding (see Chapter 14). E2F is released from the inhibiting influence of pRb in this situation and can activate its target genes.

Negative Signals

The pRb protein also serves indirectly as an integrator of negative signals emitted, in the form of a mobilization of inhibitors of CDKs. These inhibitors act on the cyclin-CDK-complexes of G_1 phase and can prevent pRb phosphorylation. The antimitogenic TGFβ, for example, strongly increases the concentration of the inhibitor $p15^{INK4b}$ via a transcription-stimulating effect. The inhibitor $p15^{INK4b}$ binds to CDK4 and CDK6 and competes with cyclin D for binding of the CDKs. This reduces the concentration of active CDK complex and hinders phosphorylation of pRb.

The tumor suppressor protein p53 (see Chapter 14) also indirectly controls the pRb function. It induces the transcription of the inhibitor p21, which inhibits the activity of CDK4/CDK6 kinase and CDK2 kinase. Because of this inhibition, activation of p53 ensures that pRb remains underphosphorylated and that the cell cycle is halted in G_1 phase. Activation of p53 is observed, for example, on damage of DNA.

Cyclin A is assigned a special role in the progess of S phase and transition into G_2 phase. Cyclin A binds and activates CDK2. The CDK2-cyclin A complex binds to the transcription factor E2F-1 and phosphorylates its DP-1 subunit. As a consequence, the DNA-binding capacity of the transcription factor is reduced and the transcription-activating function is inhibited. Furthermore, cyclin A-CDK2 complexes are involved in the phosphorylation of protein complexes involved in the initiation of DNA replication.

13.5 Cell Cycle Control of DNA Replication

Replication of DNA in S phase is subject to strict control in the cell cycle, resulting in the following observations:
– DNA replication is restricted to S phase.
– DNA is only replicated once in a cycle.
– The time sequence of DNA replication during S phase and mitosis is strictly maintained.
– If DNA damage is present, DNA replication can be stopped (DNA damage checkpoint).

Control of DNA replication occurs at two levels in particular:

Control at the Initiation Level

A main control of DNA replication occurs at the level of initiation of replication. The replication of a DNA sequence starts at specific sequence sections of the DNA, known as replication origins. In yeast, the initiation sites of DNA replication have been very well defined and characterized at the sequence level. In higher eucaryotes, in contrast, initiation has a broad initiation zone, and it has not been possible to identify a defined initiation sequence. The size of the genome in eucaryotes necessitates the use of many replication origins, which can be activated in a defined time sequence and position-specific order.

The most important components in cell cycle control of origin activity have now been identified.

Control of origin activity occurs via specific protein complexes that are bound at certain times of the cell cycle to a replication origin. For replication initiation, two states of this protein complex are important, known as the pre-replication complex and the post-replication complex (Fig. 13.17).

The *pre-replication complex* (pre-RC) is formed during anaphase and is inherited by the sister chromatids. Upon entry into S phase, the pre-RC must be disrupted for initiation to occur. If initiation begins, the pre-RC changes to the post-replication complex state, which does not permit further initiation.

Three protein complexes are central to the formation of the replication initiation complexes:

- **Origin recognition complex (ORC), composed of 6 different proteins**
  The ORC binds in a sequence-specific manner to the origin DNA and remains constitutively bound there during the whole cell cycle.
- **MCM proteins, including 6 different proteins, MCM2 – MCM7**
  The MCM proteins bind the ORC-DNA complex, but dissociate in the process of initiation during S phase. A helicase activity has been assigned to a hexameric form of the MCM complex, and it is assumed that this helicase is used for DNA unwinding at the replication fork.

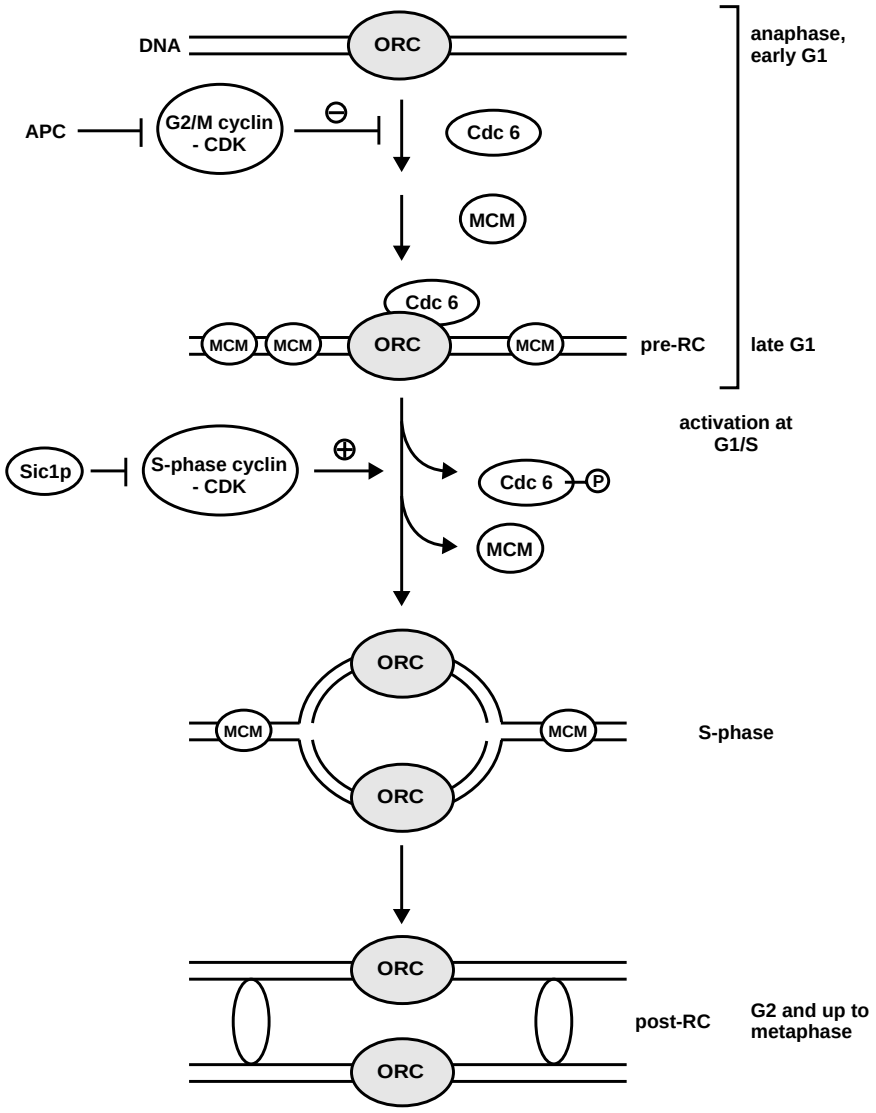


Fig. 13.17 Model of the cell cycle control of chromosomal replication in eucaryotes.
In anaphase and during G1, a prereplication complex (pre-RC) is formed at replication initiation sites which contains the Cdc6 protein and the MCM proteins, in addition to the constitutively bound origin recognition complex (ORC). Formation of pre-RC is negatively regulated by the G2/M cyclin – CDK complex. The Cdc6 protein and the MDM proteins are removed at the G1-S transition due to phosphorylation by the S-phase cyclin – CDK complex which allows the onset of DNA replication. The activity of the cyclin-CDK complexes is controlled by several means, including degradation (anaphase promoting complex, APC) and by inhibitors (Sic1 in yeast). Upon completion of S-phase, a post-replication complex (post-RC) is established. Pre-RC and post-RC cannot occur simultaneously in the cell, and this prevents re-replication of DNA in G2- and in M-phase. The cell can enter a new S-phase only after the chromatin has been licensed for a new round of DNA replication by the binding of the Cdc6 protein and MDM proteins. The diagram does not include the participation of other protein factors that have been shown to be involved in the licensing process.

- **Cdc6 protein**
  The Cdc6 protein is an essential component of the pre-RC. It is synthesized during G_2 phase and is available for formation of the pre-RC during mitosis.
- **Other proteins**
  Other proteins that license an origin for replication include the MCM10 protein and the Cdc45 protein. Furthermore, protein kinase activity is required for activation of origins (see below).

Following exit from mitosis, replication licensing occurs by recruitment of the MCM complex to replication origins. This occurs during a time window that is defined by the destruction of the mitotic regulators by the APC and the accumulation of G_1-specific initiation factors, the Cdc6 protein and the Cdt1 protein. Both proteins act synergistically to recruit the MCM complex to replication origins during G_1 phase. Activation of origins occurs upon entry into and during S phase, and this step requires the activity of two types of protein kinases, the G_1-specific CDKs with their associated cyclins and the Cdc7-Dbf4 kinase (review: Masai and Arai, 2002) . These kinases phosphorylate components of the pre-replication complex present at replication origins. The MCM2 protein has been identified as one of the substrates. Details of the order of the action of the two protein kinases and the spectrum of substrates remain to be identified.

Following activation of an origin and initiation of DNA synthesis, the Cdc6 protein becomes phosphorylated and dissociates from the origin, causing transition of the pre-replication complex to a post-replication complex that is no longer able to initiate DNA replication.

Availability of the Replication Components

At the start and during S phase, all proteins required for replication and the dNTPs must be available in sufficient quantities. An important control function is performed here by the transcription factor E2F, which induces the different enzymes needed for replication.

13.6 The G_2/M Transition and Cdc25 Phosphatase

An important decision for the entry into mitosis is taken at the G_2/M transition (review: Smits and Medema, 2001). At this point, the cell cycle can be arrested if the DNA is not properly replicated or is damaged. Therefore, the G_2/M transition functions as an important checkpoint for progression through M phase.

Entry into M phase is primarily determined by the activity of the cyclin B-CDC2 kinase complex, which is also called the mitosis-promoting factor, MPF.

Crucial regulatory elements of CDC2 activity are the concentration of cyclin B and activating phosphorylation/dephosphorylation events. The concentration of cyclin B increases with entry into S phase to a threshold at which sufficient cyclin B-CDC2 is available for triggering mitosis. Mitosis does not begin yet, however, because the CDK2-cyclin B complex requires targeted activation by CAK (see Section 13.2.1)

and dephosphorylation by members of the Cdc25 phosphatase family. A further control of cyclinB-CDC2 activity is exerted via its nuclear localization. The transport of cyclinB-CDC2 from the cytoplasm to the nucleus is also controlled by its phosphorylation status.

The regulatory phosphorylation of cyclin B-CDC2 occurs on Thr161, Thr14 and Tyr15 (Fig. 13.18). In this threefold phosphorylated form, the CDC2-cyclin B complex is inactive and remains as such until the end of G_2 phase. Targeted activation occurs at the G_2/M transition by dephosphorylation at Tyr15 and Thr14. This reaction is performed by Cdc25C phosphatase, the activity of which is controlled by phosphorylation in a way that is not yet fully established. An activated CDC2-cyclin B complex has been shown to phosphorylate and activate the Cdc25 protein providing a positive feedback loop between CDC2 and Cdc25C phosphatase. However, another upstream protein kinase is probably required for the initial activation of Cdc25C phosphatase. A candidate for this kinase is the polo-like kinase Plx1, which can phosphorylate and thus activate the Cdc25C phosphatase.

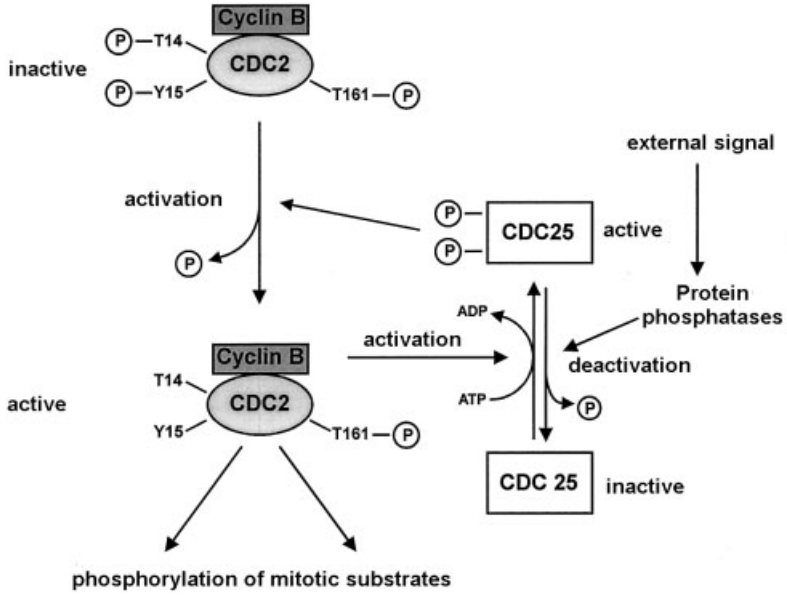


Fig. 13.18 Model of function and regulation of the Cdc25 phosphatase in *S. pombe*. The CDC2-cyclin B complex is inactive when residues Thr14 and Tyr15 are phosphorylated. Cdc25 phosphatase dephosphorylates both residues and activates the CDC2-cyclin B complex. The activity of CDC25 phosphatase is itself regulated by phosphorylation-dephosphorylation. The activated CDC2-cyclin B complex catalyzes – in addition to phosphorylation of mitotic substrate proteins – phosphorylation of the Cdc25 phosphatase. The latter is activated by the phosphorylation. A cooperative amplification of CDC2-cyclin B and a rapid transition from G_2 to S-phase take place. Other protein kinases as e. g. members of the Polo kinase family, have been implicated in phoshorylation and activation of Cdc25 too. Suppression of the activating function of CDC25 phosphatase may originate from protein phosphatases that can be activated by external signals and that dephosphorylate and inactivate Cdc25 phosphatase.

13.7 Summary of Cell Cycle Progression

The main events that control progression through the cell cycle may be summarized as follows:

- **G1 phase progression**
  Following exit from mitosis and entry into G_1 phase, cells can enter a quiescent state or they can continue in G_1, which requires the presence of mitogenic signals in the form of growth hormones. Signaling by growth hormones increases the level of D-type cyclins because of increased transcription. The increase in D-type cyclins and the formation of cyclin D-CDK4/6 complexes has at least a twofold effect. The metabolism and growth of the cells are stimulated and the cells are able to reach the critical size required for crossing of the restriction point. Furthermore, the pRb protein becomes initially phosphorylated by the cyclin D-CDK4/6 complexes, and cells are thus prepared for crossing the restriction point.
- **Activation of cyclin E/CDK2 and restriction point crossing**
  As a consequence of the increased formation of cyclin D-CDK4/6 complexes, the inhibitor $p27^{KIP1}$ is sequestered from complex formation with cyclin E-CDK2, and an initial amount of active cyclin E-CDK2 is available that continues phosphorylation of pRb and thereby initiates transcription of E2F-responsive genes, among which is the gene for cyclin E. Activation of cyclin E-CDK2 also requires active CDC25A phosphatase, which dephosphorylates the inhibitory Thr14- and Tyr15-phosphates. Now the requirements for restriction point crossing are fulfilled and the continued action of the E2F transcription factors provides for the enzymes that are necessary for entry into and progress through S phase.
- **S phase progression**
  Among the target genes of the E2F transcription factors is the gene for cyclin A, which increases at the beginning of S phase. The cyclin A-CDK2 and the cyclin E-CDK2 complexes are thought to phosphorylate important components of initiation complexes of DNA replication and thereby induce the transition of pre-replication complexes to the post-replicative state. Shortly after entry into S phase, the cyclin E is targeted for degradation in the ubiquitin-proteasome pathway, and the activity of the cyclin E-CDK2 is shut off. Further progress through S phase requires the continued action of cyclin A-CDK2 complexes.
- **G/M transition and progress through M Phase**
  During S phase and G_2 phase, the cyclin B-CDC2 complex accumulates in an inhibited state and is activated by the action of the Cdc25B/C enzymes at the G_2/M transition. The active cyclin B-CDC2 complex phosphorylates numerous substrates and is inactivated by proteolysis only at the end of M phase and during G_1 phase.

13.8 The DNA Damage Checkpoints

When the genetic material is damaged or when DNA replication is stalled, a delay in the progression of the cell cycle is initiated. Several DNA damage checkpoints exist that are able to register DNA damage and bring the cell cycle to a halt. Activation of the DNA damage checkpoints can result in arrest in either G_1 or G_2 in order to prevent replication of a damaged template. Furthermore, progress of S phase can be slowed down if the replication apparatus meets damaged DNA sites. The DNA damage checkpoints induce a series of different physiological responses including:

- halt in the cell cycle in G_1, S or G_2 phases
- slowing of DNA replication
- increased transcription of repair genes
- induction of programmed cell death, apoptosis.

The signaling pathways that lead from the appearance of DNA damage to a halt in the cell cycle involve an entire network of damage response proteins that cooperate in protecting cells from the potentially deleterious consequences of DNA damage, and this network is of the utmost importance for the prevention of cancer (review: Melo and Toczyski, 2002).

The Mammalian G_1 DNA Damage Checkpoint

This checkpoint (see Fig 13.19) serves to prevent the replication of damaged DNA (reviews: Bartek and Lukas, 2002; Rouse and Jackson, 2002; Bartek, 2003). Initiation of signaling is thought to be performed by two large protein kinases, the ATM (Ataxia telangiectasia mutated) kinase and its relative, the ATR (Ataxia and Rad-related) kinase. These kinases belong to the superfamily of PI3-like kinases (see Section 6.6.1). The two kinases function independently and are thought to be directly involved in the recognition of damaged DNA. Following activation of ATM/ATR, the signal is passed on to downstream kinases, the Chk1 and Chk2 kinases. Several other accessory proteins are involved in this step. The Chk1 and Chk2 kinases function as effector kinases that phosphorylate various downstream substrates.

Arrest in G_1 phase can now be achieved in at least two ways, depending on the substrates of the Chk1 and Chk2 enzymes. In one rapid way, the dual specificity phosphatase Cdc25C is phosphorylated on Ser 123 and is thereby targeted for ubiquitination and degradation in the proteasome pathway. The lack of this enzyme locks the CDK2 kinase in the inactive form phosphorylated on threonine 14 and tyrosine 15. The cyclin E-CDK2 complex that is required for entry into S phase is inhibited, and the cell cycle arrests at G_1/S. It should be noted that the scheme in Fig. 13.19 is only a minimal scheme that does not address the participation of numerous other proteins that function as adaptors or structural proteins in these processes.

In a second way, which appears to sustain a prolonged arrest in G_1, the tumor suppressor p53 becomes phosphorylated by Chk1/Chk2 on serine 20, leading to stabilization and accumulation of p53. As a consequence, the transcription of the CDK inhibitor $p21^{Cip1}$ is increased and the $p21^{Cip1}$ now can inhibit all CDK complexes.

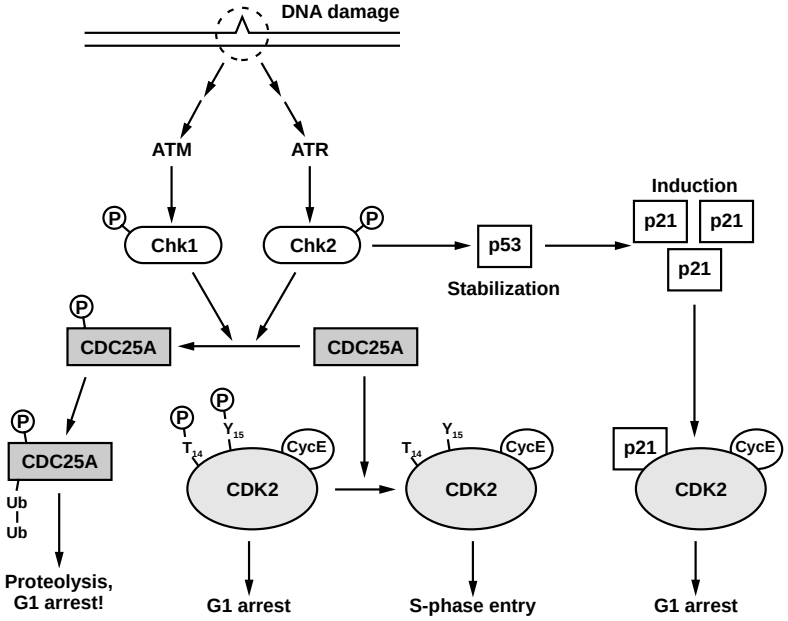


Fig. 13.19 The G_1 DNA damage checkpoint. DNA damage induces activation of the protein kinases ATM and ATR via reactions not illustrated in the figure. ATM and ATR phosphorylate and activate the protein kinases Chk1 and Chk2 which phosphorylate and inactivate the protein phosphatase CDC25A. This phosphatase is required for removal of the inhibitory phosphorylations on T14 and Y15 of CDK2 and hence for G_1/S-phase progression. CDC25A phosphorylation mediates ubiquitination and proteasomal degradation of CDC25A. As a consequence, CDK2 retains its inhibitory phosphorylations and the cell arrests in G1.

The G_2 DNA Damage Checkpoint

The G_2 damage checkpoint uses many of the components of the G_1 damage checkpoint. Following damage sensing and activation of ATM/ATR, the Chk1/Chk2 kinases are activated and phosphorylate the phosphatase Cdc25B/C (see above) on distinct serine residues. The phosphoserine residues now serve as attachment points for members of the 14-3-3 protein family. Binding of 14-3-3 proteins can either directly inhibit the phosphatase or influence the subcellular distribution of Cdc25B/C, leading to cytoplasmic export and separation from cyclinB-CDC2. The phosphatase activity is no longer available for activation of cyclinB-CDC2, and entry into M phase is blocked.

Reference

1 Bartek, J. and Lukas, J. (2001) Mammalian G1- and S-phase checkpoints in response to DNA damage. *Curr.Opin.Cell Biol.*, **13**, 738–747.

2 Bartek, J. and Lukas, J. (2003) DNA repair: Damage alert. *Nature*, **421**, 486–488.

3 Dhavan, R. and Tsai, L.H. (2001) A decade of CDK5. *Nat.Rev.Mol.Cell Biol.*, **2**, 749–759.

4 Endicott, J.A., Noble, M.E. and Tucker, J.A. (1999) Cyclin-dependent kinases: inhibition and substrate recognition. *Curr.Opin.Struct.Biol.*, **9**, 738–744.

5 Gieffers, C., Dube, P., Harris, J.R., Stark, H. and Peters, J.M. (2001) Three-dimensional structure of the anaphase-promoting complex. *Mol.Cell*, **7**, 907–913.

6 Harper, J.W., Burton, J.L. and Solomon, M.J. (2002) The anaphase-promoting complex: it's not just for mitosis any more. *Genes Dev.*, **16**, 2179–2206.

7 Kaldis, P. (1999) The cdk-activating kinase (CAK): from yeast to mammals. *Cell Mol.Life Sci.*, **55**, 284–296.

8 Kasten, M. and Giordano, A. (2001) Cdk10, a Cdc2-related kinase, associates with the Ets2 transcription factor and modulates its transactivation activity. *Oncogene*, **20**, 1832–1838.

9 Lee, M.H. and Yang, H.Y. (2001) Negative regulators of cyclin-dependent kinases and their roles in cancers. *Cell Mol.Life Sci.*, **58**, 1907–1922.

10 Masai, H. and Arai, K. (2002) Cdc7 kinase complex: a key regulator in the initiation of DNA replication. *J.Cell Physiol*, **190**, 287–296.

11 Michael, D. and Oren, M. (2002) The p53 and Mdm2 families in cancer. *Curr.Opin.Genet.Dev.*, **12**, 53–59.

12 Obaya, A.J. and Sedivy, J.M. (2002) Regulation of cyclin-Cdk activity in mammalian cells. *Cell Mol.Life Sci.*, **59**, 126–142.

13 Page, A.M. and Hieter, P. (1999) The anaphase-promoting complex: new subunits and regulators. *Annu.Rev.Biochem.*, **68**, 583–609.

14 Pavletich, N.P. (1999) Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. *J.Mol.Biol.*, **287**, 821–828.

15 Peters, J.M. (1998) SCF and APC: the Yin and Yang of cell cycle regulated proteolysis. *Curr.Opin.Cell Biol.*, **10**, 759–768.

16 Rouse, J. and Jackson, S.P. (2002) Interfaces between the detection, signaling, and repair of DNA damage. *Science*, **297**, 547–551.

17 Russo, A.A., Jeffrey, P.D., Patten, A.K., Massague, J. and Pavletich, N.P. (1996) Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. *Nature*, **382**, 325–331.

18 Russo, A.A., Tong, L., Lee, J.O., Jeffrey, P.D. and Pavletich, N.P. (1998) Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature*, **395**, 237–243.

19 Sano, M. and Schneider, M.D. (2003) Cyclins that don't cycle: cyclin t/cyclin-dependent kinase-9 determines cardiac muscle cell size. *Cell Cycle*, **2**, 99–104.

20 Simone, C. and Giordano, A. (2001) New insight in cdk9 function: from Tat to MyoD. *Front Biosci.*, **6**, D1073-D1082.

21 Smits, V.A. and Medema, R.H. (2001) Checking out the G(2)/M transition. *Biochim.Biophys.Acta*, **1519**, 1–12.

22 Stiegler, P. and Giordano, A. (2001) The family of retinoblastoma proteins. *Crit Rev.Eukaryot.Gene Expr.*, **11**, 59–76.

23 Tomoda, K., Kubota, Y. and Kato, J. (1999) Degradation of the cyclin-dependent-kinase inhibitor p27Kip1 is instigated by Jab1. *Nature*, **398**, 160–165.

24 Trimarchi, J.M. and Lees, J.A. (2002) Sibling rivalry in the E2F family. *Nat.Rev.Mol.Cell Biol.*, **3**, 11–20.

25 Tyers, M. and Jorgensen, P. (2000) Proteolysis and the cell cycle: with this RING I do thee destroy. *Curr.Opin.Genet.Dev.*, **10**, 54–64.

26 Weinberg, R.A. (1995) The retinoblastoma protein and cell cycle control. *Cell*, **81**, 323–330.

27 Yew, P.R. (2001) Ubiquitin-mediated proteolysis of vertebrate G1- and S-phase regulators. *J.Cell Physiol*, **187**, 1–10.

14
Malfunction of Signaling Pathways and Tumorigenesis: Oncogenes and Tumor Suppressor Genes

14.1
General Aspects of Tumor Formation

14.1.1
Characteristics of Tumor Cells

Tumor cells have special features compared to normal cells. The phenotype of a tumor cell is characterized by the following characteristics:
– increased rate of cell division, loss of normal growth control
– loss of ability to differentiate
– loss of contact inhibition
– increased capability for invasion of neighboring tissue, metastasis.

The cells of a fully grown, aggressive tumor have acquired these properties in a slow, multi-step process with the characteristics of *cellular evolution*. This development is associated with a *selection process*, in the course of which, cells that have lost their growth-regulating mechanisms predominate. The transition of a normal cell to a tumor cell can often be well characterized both morphologically and physiologically. However, only a few of the underlying biochemical changes are understood. Tumor cells differ from their progenitor cells in having acquired a large number of genetic changes. It is estimated that tumor cells accumulate several thousand to several hundred thousand changes in DNA sequence.

The processes that initiate, contribute to and propagate the malignant phenotype can be summarized as follows (Fig. 14.1).

- **Initial mutations in repair enzymes and/or cell cycle checkpoint genes**
  The mutation and malfunction of key components of DNA repair and DNA damage checkpoints will lead to an increased mutational load on the cell and will favor the accumulation of further genetic changes. One hypothesis is that tumor cells acquire – at some state of the tumor development – a mutator phenotype, which is characterized by an increased mutation rate due to a malfunction of gene products responsible for the maintenance of genetic stability (see Section 14.1.5).

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

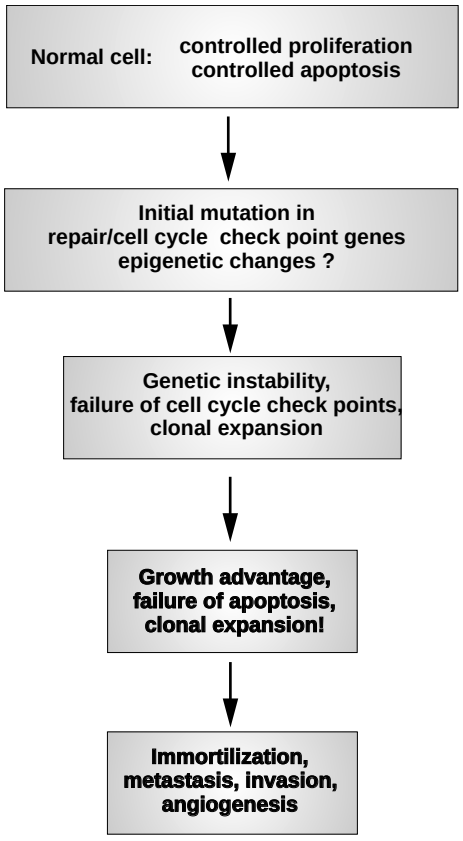


Fig. 14.1 Model of the transition of a normal cell into a tumor cell.

- **Enhanced genetic instability**
  Changes in gross chromosomal structure and instability at the level of DNA repeats (microsatellite instability) are characteristics of many tumor cells. These alterations arise from the malfunction of DNA repair and the checkpoints that couple DNA damage to cell cycle arrest and apoptosis.
- **Growth advantage, selection and expansion of tumor cell pools**
  Increased mutation rates and enhanced genetic instability will create a pool of tumor cells that contain a large number of variants. At this stage, the tumor is only clonal in the sense that it was originally derived from a single progenitor stem cell. Rather, the tumor must be considered as a population under change, harboring a huge collection of coexisting subclones. This heterogeneous and dynamic population is the target of selection for cells that have a growth advantage due to escape of the normal cellular control of cell proliferation and are no longer subject to the normal cell death mechanisms. Selection will allow for the outgrowth of those cells which can proliferate and grow optimally in the surroundings of the progenitor cell. As a consequence, the pool of tumor cell variants will expand.

- **Formation of solid tumors and metastasis**
  The pool of tumor cell variants must be considered as a highly flexible collection of different cells with the potential for future changes in the presence of selective pressures. Eventually, subclones will appear that have the ability to proliferate independently of signals of the neighboring cells and have lost contact inhibition. Thus, in a late stage of tumorigenesis, tumor cells can acquire the potential to survive in a foreign cellular environment and form organ-like structures. Cells of the final tumor carry selected alterations, e. g., altered metabolism and the ability to form blood vessels, along with the instability mutations.

Although most of the processes contributing to tumorigenesis have now been identified, there is still controversy concerning the timing and the nature of the initial tumor-causing events. Whereas there is general agreement that tumor formation is triggered by the accumulation of several genetic changes, the exact nature of the initial genetic changes in tumor formation is still a matter of discussion. Additionally, it is becoming increasingly clear that the difference between normal cells and tumor cells has to be described not only in terms of alterations in DNA sequence and chromosome structure, but also in terms of changes in DNA modification, namely DNA methylation. The latter point refers to epigenetic events that influence the protein profile of a cell, not at the level of DNA sequence but by alterations in the DNA methylation pattern.

14.1.2 Genetic Changes in Tumor Cells

A general phenomenon in tumor formation is the *stepwise accumulation* of changes in genetic information (mutations). DNA sequence analysis and gene expression profiling have shown that cancer cells differ from the progenitor cells in the large number of genetic changes. These genetic changes are passed on from the mother cell to the daughter cell during cell division. The mutations observed in tumors include a broad spectrum of reorganizations and changes in genetic information. The extent of the changes can be very varied. Smaller mutations are observed such as
– simple base substitutions
– insertion or deletion of bases
– inversions and duplications of DNA sequences.

The mutations can lead to changes in genes, with products involved in DNA repair and DNA damage control, in signal transduction pathways, and in growth regulation. It is assumed that these mutations are particularly important in the early phase of tumor formation.

In the later phases of tumor formation, an increasing *genetic instability* occurs, which is visible at the level of the chromosomes, particularly in the form of a change in the normal chromosome number (review: Lengauer et al., 1998).

The following are observed:
– loss or duplication of whole chromosomes
– multiplication of the chromosome set

– chromosome translocations: deletion, addition or exchange of individual chromosomes
– amplification of DNA sequences.

Changes in the chromosome structure are often observed in tumors of the blood-forming system, the leukemias and lymphomas. They are almost always found in the later phases of aggressive solid tumors. These extensive reorganizations have far-reaching consequences for growth behavior and functional performance.

14.1.3 Epigenetic Changes in Tumor Cells

In addition to the direct changes in the information content of the DNA, epigenetic events are now known to contribute significantly to tumor formation (review: Jones and Baylin, 2002). Epigenetics is defined as heritable modifications of the genome that are not accompanied by changes in DNA sequence. Most epigenetic phenomena are linked to DNA methylation. It has been outlined already in Section 1.4.8 that methylation at CpG sequences is a DNA modification that is heritable during cell division. This is a major tool for controlling gene expression operating via changes in chromatin structure.

Epigenetic alterations in cancer cells include global hypomethylation, hypomethylation of individual genes, and hypermethylation of CpG-rich sequences. Almost half of the mammalian promotors contain CpG-rich sequences that are unmethylated in normal cells. Hypermethylation of these sequences has now been recognized to be the most prominent epigenetic event in tumor cells, and it is found in virtually all human tumors. Such promotor hypermethylation will lead to inappropriate gene silencing, and the effect may be the same as mutation of the gene. Many genes linked to tumor formation have been found to contain changes in the methylation pattern, and there is now a growing list of tumor suppressor genes (see Section 14.4) that are silenced by promotor hypermethylation in certain cancers. In many tumors, loss of function of tumor suppressor genes can only be explained by hypermethylation of the promotors, since no changes in DNA sequence could be found in these genes. Other genes whose function is disrupted by aberrant DNA methylation in tumors include repair enzymes and cell cycle regulators.

Table 14.1 gives examples of hypermethylation of genes linked to tumor formation.

It should be noted that methylation-linked repression of repair genes like that for O6-methylguanine-DNA-methyltransferase will impair DNA repair and will increase mutation rates. Furthermore, methylated CpG sequences are more prone to mutagenesis, because spontaneous deamination of 5-methyl-C will induce a C→T transition. Gene silencing by hypermethylation may therefore favor the occurrence of mutational events during tumorigenesis.

Tab. 14.1 Genes that are frequently mutated and/or hypermethylated in cancer. Only a subset of genes that are known to be frequently silenced in one or more cancer types are shown.

Gene	Function
p14, p15	CDK inhibitors
PTEN	$PtdInsP_3$ phosphatase
MLH1	DNA mismatch repair
APC	Wnt signaling
CDH-1	Cell adhesion, loss during tumor metastasis
NF1	GAP involved in Ras signaling
TIMP3	Tissue inhibitor of metalloproteinase 3
BRCA1	DNA repair
STK11	Serine/threonine kinase 11, mutated in hereditory cancers
MGMT	DNA repair

Abbreviations and links: : **p14, p15**, CDK inhibitors of Ink family (see section 13.2.6); **PTEN** see section 6.6.3; **MLH1**, mutL homolog 1, see section 14.5; **APC**, adenomatosis polyposis coli, see section 14.9; CDH1, E-cadherin cell adhesion protein, see section 14.9; **NF1**, neurofibromin 1, GTPase activating protein, see section 9.5; **TIMP3**, tissue inhibitor of metalloproteinase 3; **BRCA1**, breast cancer 1, early onset, see section 14.5; **STK11**, serine/threonine kinase 11; **MGMT**, O^6-methylguanine-DNA methyltransferase.

14.1.4 Causes of Oncogenic Mutations

Mutations that trigger and promote tumor formation may be caused by a number of processes. Briefly summarized, the following factors are involved:

1. **Intrinsic changes in the genetic information**
   There are many endogenous processes that change the information content of the DNA. These processes are inseparable from the performance and characteristics of an organism, and they are influenced by external factors to only a small extent.
   DNA replication errors
   Replication of DNA does not occur with complete precision, but rather has an intrinsic inaccuracy. The error rate for incorporation of nucleotides in DNA replication is of the order of one error per $10^6 - 10^8$ correctly incorporated nucleotides.
   Spontaneous DNA damage
   DNA has only limited stability in the temperature and pH conditions of an organism. Spontaneous changes in the DNA structure may occur, and cleavage of purine bases is considered to have an important role. The apurinic sites resulting from depurination may give rise to mutations if not repaired.
   Metabolism-related damage of DNA
   Reactive metabolic products are an important cause of endogenous damage to DNA. In the first place is reactive oxygen, which can lead to DNA damage, especially in the form of the superoxide anion O_2^- and in the form of OH radicals. Dietary components also include many chemical compounds with which DNA

can react. Furthermore, dietary components can be converted by metabolic activation into compounds with a high potential for damage to DNA.

2. **External damage of DNA**
   A number of external factors can lead to DNA damage and thus to mutations. These include especially the effects of carcinogenic chemicals and UV, X-ray, and other high-energy radiation (e. g., radioactivity).
3. **Viruses**
   There are a large number of RNA- or DNA-viruses that are directly associated with tumor formation. Three principal mechanisms are responsible (see Section 14.3):
   – introduction of viral oncogenes into the host genome
   – interaction with virus-specific proteins and signal proteins of the host cell
   – control of expression of proto-oncogenes by a viral promoter.

14.1.5 DNA Repair, DNA Damage Checkpoints, and Tumor Formation

Because of the many damaging influences on the DNA, it is essential for the cell that DNA damage can be repaired. The repair systems involved, such as the enzyme apparatus of excision repair, the alkyl transferases, and the mispairing repair system, operate almost without error and can repair the vast majority of DNA damage. A deficit in the repair capacity of a cell, for example, due to inactivating mutation of a repair component, favors the accumulation of mutations and gives tumor formation a boost. Proteins that are involved in the avoidance of DNA damage, either directly or indirectly, make an important contribution to the suppression of tumors. Thus, mutated tumor suppressor proteins are often found in association with a hereditary predisposition towards tumor formation (hereditary tumor syndrome), and many proteins involved in the repair of DNA damage are tumor suppressor proteins (see Section 14.5). For example, a hereditary deficit in nucleotide excision repair is associated with increased formation of skin tumors. Cancer cells harbor a large number of genetic changes, and it can be estimated that these changes are not caused by the normal, natural mutation rates. Rather, an elevated mutation rate has to be assumed – and has also been demonstrated in some cases – to explain the large number of genetic changes in tumors (review: Jackson and Loeb, 2001).

These observations have led to the theory that, in an early stage of tumor formation, a mutation occurs in a repair system needed to maintain the integrity of the genome (Loeb, 1991; Loeb, 2001). Loss of the function of the repair system leads to a *mutator phenotype*: a missing or ineffective DNA repair favors further accumulation of mutations and leads to an intrinsic instability of the genome.

Of course, this need not be true for all cancers. To avoid mutations, it is equally important to have a functional DNA checkpoint system that stops cell cycle progression in the presence of DNA damage. The cell cycle arrest can allow for repair of the damaged site or it can be a signal for the cell to die by apoptosis. Accordingly, gene products that are involved in DNA damage checkpoints and/or link DNA damage to apoptosis are generally found to be functionally inactivated in tumor cells.

An enhanced mutation rate as a consequence of failure of DNA repair or DNA damage checkpoints will expand the reservoir of pre-cancerous cells that are the target of selection during tumorigenesis. The generation of a large pool of mutant cells will favor the outgrowth of cells that have lost normal cell cycle control.

Enhanced genetic instability is a prevalent property that distinguishes tumor cells from normal cells. Two main forms of genetic instability have been identified in tumor cells: microsatellite instability and chromosomal instability. Microsatellite DNA contains a large number of repeat sequences that undergo changes in both repeat sequence and number following a defect in DNA mismatch repair (see below). Genetic instability due to breakdown of a repair component is only seen in a small fraction of tumors, however. In the majority of tumors, instability at the level of the chromosomes can be established, which is manifested in loss or gain of whole chromosomes or large parts thereof. The reasons for this instability are largely unknown and are thought to reside in defects in mitotic processes or in DNA repair or DNA replication checkpoints. We know of only a few candidate proteins involved in this dysregulation. Examples are the Bub protein, which is part of a metaphase-anaphase checkpoint and the tumor suppressor APC (see Section 14.9).

14.1.6
Cell Division and Tumor Formation

Cell division itself is also assigned a central function in tumor formation. Pre-cancerous damage to the DNA, such as the formation of an adduct of a base and a carcinogenic compound, can then only lead to mutations if the cell undergoes a round of division with its damaged DNA. On replication of the damaged DNA, there is an increased probability of changes in the DNA sequence unless the error is repaired beforehand. These changes will then be passed on to subsequent generations as mutations.

Furthermore, each cell division round has the potential danger of rearrangement of chromosome sections during mitosis and thus chromosome aberrations. A thorough theoretical and experimental analysis of the dose-effect relation of various carcinogenic substances has shown that an increased cell division activity is an important risk factor for the creation of tumors (Cohen and Ellwein, 1990). All processes that lead to an increase in the rate of cell division will increase the probability of tumor formation, according to these investigations.

14.2
Cell Division Activity, Errors in Function of Signal Proteins, and Tumor Formation

DNA damage and the resulting mutations are generally randomly distributed over the genome and thus affect all genes of an organism with equal probability. Within the confines of the cellular evolution process leading from the normal cell to a tumor cell, cells with normal growth control of the organism removed by mutation are favored.

Mutations of genes involved in growth control of a cell are therefore of particular importance in tumor creation.

The preceding chapters have shown that growth and cell division are controlled by a complicated process involving many protein complexes and in which intrinsic and external control mechanisms are effective. The process includes mitogenic signals, such as growth factors, and antimitogenic signals, such as TGFβ (see Sections 12.1, 13.1.3, and 13.2.6). The external signals serve to adapt cell division activity of a cell to the function of the organism.

14.2.1 The Fate of a Cell: Quiescence, Division, or Death

A cell can enter various stages depending on the intensity and balance of mitogenic and antimitogenic signals (Fig. 14.2):

- **Continuous cell division**
  Continuous cell division is found, for example, in stem cells, which serve as precursors for other cells, and in tissue in which dying cells must be replaced. This requires the continuous effects of mitogenic signals. In this case, cell division must

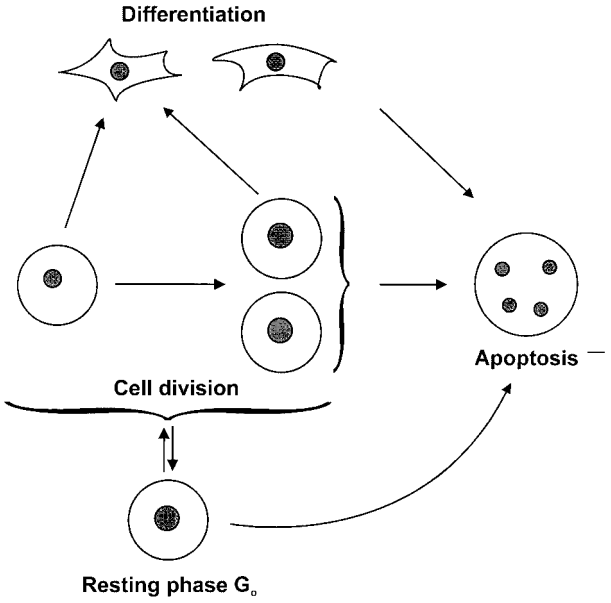


Fig. 14.2 Division activity and fate of a cell. Cells can move from a division-active state to a differentiated state or into the resting phase Go. The transition into Go phase is generally reversible, whereas differentiated cells often cannot return to the division-active state. The cell can be directed from all three stadia into programmed cell death, or apoptosis.

ensure homeostasis of the cell number or the tissue. An increase in the cell number through cell division compensates for loss of cells that die as a part of normal cell turnover or that are eliminated by programmed cell death, or apoptosis.

- **Transition into a quiescent phase**
  In the absence of external mitogenic signals or in the presence of a majority of antimitogenic signals, cell division activity may be stopped. The cell enters the quiescent phase (G_o phase). From G_o phase, cell division may be resumed when mitogenic signals reappear.
- **Differentiation of cells**
  During the development of an organism, many cells enter a state of terminal differentiation, in which they perform a specialized function. Differentiated cells originate from dividing stem cells, and, in the process, partially or completely lose the ability to divide. They are then no longer able to receive and act upon mitogenic signals.
- **Apoptosis**
  Another route that a cell can take is the pathway of programmed cell death, or apoptosis (see Chapter 15). Many different factors may be involved in triggering apoptosis, some of which are also components of mitogenic or antimitogenic signaling pathways.

14.3 Definition and General Function of Oncogenes and Tumor Suppressor Genes

A property commonly ascribed to tumor cells is enhanced cell proliferation. Yet, tumors most often originate from differentiated, adult cells that have mostly lost the ability to divide. In these cases, tumor cells arise from progenitor cells that have regained the capability for proliferation and can bypass quiescence. Dedifferentiation and altered response to mitogenic signaling are characteristics that are observed in the early stages of tumor development.

The proliferative advantage of tumor cells can be due to increased mitogenic signaling and/or lowering of the threshold required for transition from the quiescent to the dividing state. As outlined in Chapter 13, mitogenic and antimitogenic signals are received in the cell cycle controlling cell cycle entrance and exit. Most of these signals become active at the G_0 to G_1 transition by regulating the activities of CDKs, cyclins, and CKIs required for G_1 progression.

The mitogenic signals that the organism uses to control division activity of individual cells are registered by corresponding receptors and converted into intracellular signal chains that finally influence the activity of the central cell cycle regulators. As outlined in the preceding chapters, many proteins participate in these signal transduction pathways, and it is not surprising that a large number of genes have been identified, which, when mutated or not functioning normally, lead to a fault in growth control and correlate with tumor formation. The mutated genes are roughly divided into *oncogenes* and *tumor suppressor genes*. This division is based on the effect of a mutation on function.

14.3.1 Oncogenes and Proto-Oncogenes

The term oncogene was coined in association with the search for the tumor-causing principle in retroviruses. Retroviruses contain RNA as the genetic material and can transcribe RNA into DNA with the help of the virus's own enzyme reverse transcriptase. The DNA form of retroviruses can integrate into the DNA of the host cell and, during cell division, is passed on to the daughter cells as a provirus. From the provirus, viral RNA and complete virus particles may be formed.

Some representatives of the retroviruses cause tumors in animals such as mice or chickens. The discovery of oncogenes followed the investigation of the *src gene* of *Rous sarcoma virus*, which could be identified as the tumor-causing principle of this retrovirus. The *src* gene codes for the Src tyrosine kinase (see Section 8.3.2). The gene sections of retroviruses responsible for tumor formation were designated *oncogenes*.

Soon after discovery of the oncogene, it was demonstrated that the viral oncogenes are mutated forms of the genes of cellular proteins that are also active in normal cells. The cellular variants of viral oncogenes were named *proto-oncogenes*.

The oncogenes of retroviruses are prefixed with a *v* (e. g., *v-src, v-sis*), whereas the corresponding proto-oncogenes are prefixed with a *c* (e. g., *c-src, c-sis*).

Further investigations revealed that proto-oncogenes can also be converted, via activating mutations, into oncogenes, without involvement of viruses. The use of the term "oncogene" was thus extended and its definition was made more general.

Oncogenes are genes that can result in a transforming or immortalizing phenotype on experimental transformation in cellular model systems. Oncogenes arise by activating mutation of their precursors, the proto-oncogenes. Proto-oncogenes are often directly involved in growth regulation of normal cells. Oncogenes generally have *dominant* character. The mutation of a proto-oncogene to an oncogene is phenotypically visible when only one of the two copies of the gene in a diploid chromosome set is affected by the mutation. The dominant mutation is accompanied by a "*gain of function*"; it typically amplifies or increases the yield of a function in growth regulation.

Generally the function of oncogenes during multistep carcinogenesis is considered to reside in a growth advantage that is conferred by the altered function of the proto-oncogene product. The affected cells are no longer subject to the normal growth control that is exerted by a finely balanced network of mitogenic and antimitogenic signals. Therefore, cells harboring oncogenes proliferate preferentially and autonomously and help to expand the pool of mutated cells that is the target of selection during tumorigenesis (see Section 14.1).

It should be pointed out that the oncogenic function of a particular gene is generally defined via cellular model systems which use primary cell lines or immortalized cell lines. The value of these experiments is highlighted by the observation that many of the transforming or immortalizing genes in cellular model systems are also found in a similarly mutated form in tumors in man or mice. The systems, however, do not mirror the situation in a real tumor. Rather, only limited and selected aspects of tumorigenesis can be studied. Importantly, the multistep nature of tumor formation is not considered in the cellular model systems.

14.3.2
Mechanisms of Activation of Proto-Oncogenes

The activation of a proto-oncogene to an oncogene is based on mutations that can change the function and regulation of the affected protein by various mechanisms. Two pathways of activation can be roughly differentiated (Fig. 14.3). On the one hand, the structure of the coded protein may be affected; on the other hand, activation may lead to a concentration increase in the protein.

Activation by Structural Changes

A frequent cause of activation of proto-oncogenes is a change in the structure of the coded protein, affecting the regulation and function. Via the oncogenic activation, there is no creation of completely new functions, but rather the normal function of a proto-oncogene product is modified and/or released from cellular regulation.

The spectrum of structural mutations that can convert a proto-oncogene into an oncogene is very diverse. Both simple amino acid changes and larger structural

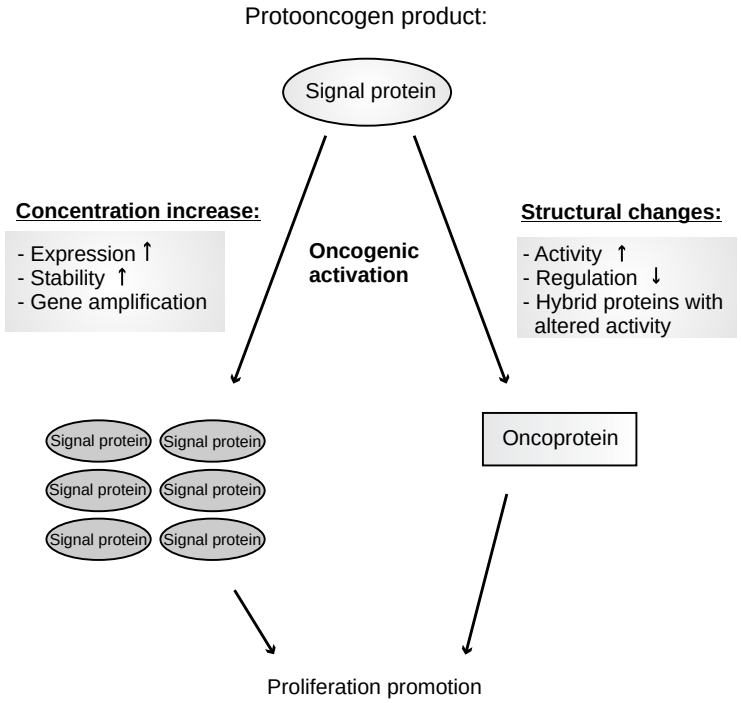


Fig. 14.3 Mechanism of activation of proto-oncogenes to oncogenes. Proto-oncogenes may be converted into oncogenes via the concentration increase pathway or the structural change pathway. In the case of the concentration increase, there is an excessive and unprogrammed function of the signal protein coded by the proto-oncogene. In the case of structural change, the proliferation-promoting activity of the oncoprotein results from changed activity, altered regulation or formation of a hybrid protein.

changes are observed in the coded protein. In particular, viral oncoproteins demonstrate multiple mutations compared to their cellular counterparts, linked to important and far-reaching structural and functional changes.

Increase in Activity of Signal Proteins

Activating point mutations may directly affect the enzyme activity of a signal protein. The mechanism of activating oncogenic point mutations is best documented for the Ras protein. Many of the point mutations at positions 12 and 13 of the Ras protein, which are also observed in solid tumors, bring about a reduction in the rate of GTP hydrolysis. This can no longer be increased by the GAP protein, so that the active GTP state exists for a longer duration (cf. Section 9.2.3). It is assumed that this is the cause of unprogrammed stimulation of the Ras-MAPK pathway, which is manifested as increased cell division activity and leads to a growth advantage of the affected cell.

Change in the Regulation of Signal Proteins

Mutations may lead to loss of cellular control over the activity of a proto-oncogene. Frequently, this brings about constitutive activation of the signal protein. Thus, in the transforming *v-raf* gene, the N-terminal sequence section of Raf kinase is missing, on which both the autoinhibitory function and the phosphorylation sites of Raf kinase are localized (see Section 9.6).

Formation of Hybrid Proteins

In many tumors, a reciprocal exchange of DNA sections on different chromosomes is observed. During this translocation of chromosomes, gene fusions may occur, leading to the formation of chimeric proteins. Within the chimeric proteins, there are often structural portions that originate from signal proteins. The function of the signal protein portion is removed from normal regulation in the chimeric protein and can have a tumor-promoting effect. Tyrosine kinases and transcription factors are often affected by gene fusions. The chimeric proteins arising from chromosome translocation frequently represent a characteristic of a particular tumor type. A review of gene fusions observed as a consequence of chromosome translocations in tumors is given in Look (1997).

Activation by Concentration Increase

A change in the gene expression or stability of a proto-oncogene product may lead to an increase in the cellular concentration of the protein. Because of the increased concentration, a mitogenic signal mediated by a proto-oncogene product may be amplified.

Overexpression of Proto-oncogenes

The overexpression of proliferation-regulating proteins may lead to immortalization and/or transformation in cellular model systems. Overexpression of signal proteins is observed in many tumors, and it is assumed that the *overexpression is associated with pathogenesis of the tumor*. The mechanisms leading to overexpression are diverse and theoretically include all processes involved in expression regulation. Of note are translocations of a proto-oncogene into the vicinity of a strong promoter, causing excessive

or deregulated expression of the corresponding gene. Stabilization of the mRNA of a proto-oncogene by deletions at the 3' end can also lead to overexpression, as has been shown for cyclin D1.

Activation of proto-oncogenes by unprogrammed expression is often associated with chromosome translocations in leukemias and lymphomas. In Burkitt's lymphoma, different translocations of the *c-myc* gene, which codes for the c-Myc transcription factor, are found. The translocations bring about movement of the *c-myc* gene into the vicinity of immunoglobulin genes. Consequently, constitutive expression of the c-Myc protein occurs, disturbing the normal regulatory network into which the c-Myc protein is bound (see Section 14.3.3).

Another tumor-promoting mechanism based on deregulation of transcription of growth factors is the formation of *autocrine loops*. In the course of tumor formation, unprogrammed expression of growth factors may occur in cells in which there would normally be little or no expression of these proteins. If these cells express the appropriate growth factor receptors, the growth factors may bind to these and create a stimulus of division. The cell is no longer dependent on the supply of an external growth factor. The cell then produces its own growth factor and division stimulus (Fig. 14.4).

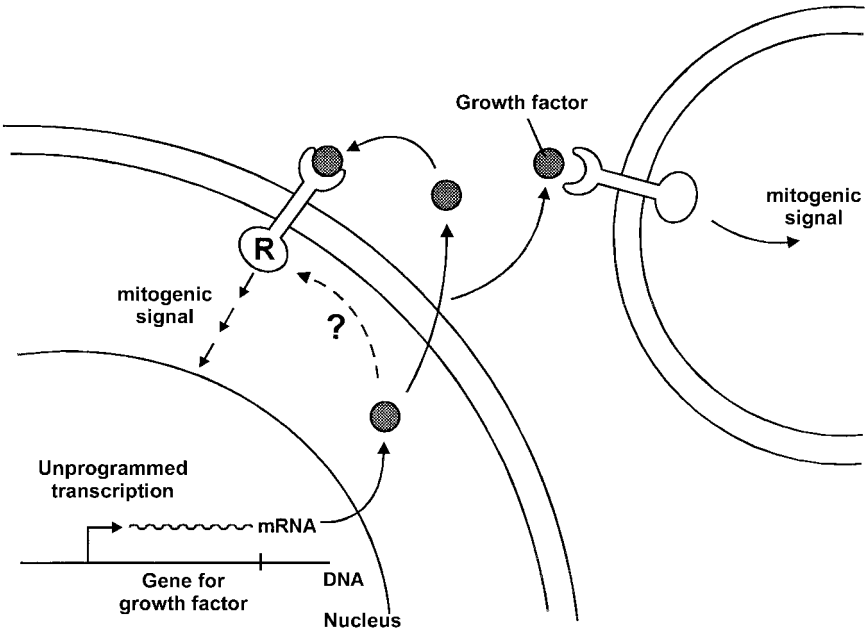


Fig. 14.4 Autocrine loops in tumor formation. Due to an error in control of transcription, growth factors may be produced and secreted in the cell which would normally only be formed in low concentrations or not at all. If the cell also possesses the receptors corresponding to the growth factor, the growth factor can then bind and activate a mitogenic signal chain. In this situation, the cell creates the mitogenic signal itself. There is evidence that the growth factors can become active intracellularly. The mechanism behind this is unknown.

Amplification of Genes for Signal Proteins

Another factor closely associated with tumor formation is an increase in the number of copies of a gene for a signal protein (*gene dose effect*). In a number of tumors, amplification of genes with products that have a central position in growth regulation is observed. The tumor-promoting effect of amplification of genes coding for signal proteins can also be observed in cellular model systems.

Increased Stability of Proto-oncogenes

The concentration of many signal proteins is controlled via ubiquitin-mediated proteolysis. It is therefore not surprising that proteins have been discovered that, when mutated, increase the lifetime of signaling proteins and thereby have an oncogenic effect. One example is the Cbl-protein, which functions as an E3 enzyme in the degradation of cell surface receptors. Another example of an oncogene not directly involved in signal transduction is the Bcl-2 protein, which is involved in apoptosis (see Section 15.4). Oncogenic mutation of Bcl-2 correlates with decreased apoptosis and allows the survival of cells that normally would be eliminated by the apoptotic program.

14.3.3 Examples of the Functions of Oncogenes

More than 100 dominant oncogenes have been identified to date. Nearly all the components of the signal transduction chains that transmit signals from the cell exterior to the level of the cell cycle and transcription can be converted by mutations into an oncogenic state. It has to be emphasized that activation of an oncogene is generally not sufficient for transformation of a normal cell into a tumor cell. As outlined in Section 14.8.6, functional inactivation of cell death mechanisms must occur at the same time, and it is the cooperation of defects in oncogenic pathways and in apoptotic pathways that paves the way for the formation of a real tumor.

Only the most important and best-characterized oncogenic signaling proteins are discussed below.

Receptor Tyrosine Kinases

More than the half of the known receptor tyrosine kinases have been repeatedly found in either mutated or overexpressed forms in human malignancies, including sporadic cancers. Most activating mutations lead to a ligand-independent, constitutive activation of the tyrosine kinase activity of the receptor. As outlined in Chapter 8, receptor tyrosine kinases normally exist in a repressed state and require ligand-induced autophosphorylation for activation. Oncogenic receptor tyrosine kinases often escape from this control and induce inappropriate activation of downstream signaling components that leads to enhanced cell proliferation and increased cell survival (review: Blume-Jensen and Hunter, 2001).

A well-studied example is provided by the ErbB2 (also called Her or Neu) and EGF receptors, which are receptor tyrosine kinases that form heterodimers with each other and with two other RTKs, ErbB3 and ErbB4, which have various external growth fac-

tors as ligands. No external ligand is known for ErbB2, and heterodimerization with the three other receptors is required for ErbB2 signaling. Overexpression of ErbB2 in cancer cells disturbs the complex network of receptor interactions and induces the formation of ErbB2-containing heterodimers. These have a potent oncogenic effect due to inappropriate activation of mitogenic and survival pathways such as the MAPK pathways and the PI3 kinase pathway.

Other oncogenic mechanisms of RTK activation include mutations in the kinase domain inducing conformational shift of the activation loop and a subsequent constitutive activation as well as constitutive dimerization as the result of a mutation in the extracellular domain of the RTK. Both modes of activation have been observed in germline mutations of the receptor tyrosine kinase Ret, leading to heritable tumor syndromes.

Nonreceptor Tyrosine Kinases

Many of the nonreceptor tyrosine kinases were discovered because the mutated form of the protein is the product of a viral oncogene. The most prominent examples are the Src tyrosine kinase and the Abl tyrosine kinase (see Section 8.3). The relationship of the Abl tyrosine kinase with the Philadelphia chromosome translocation in lymphocytes has been especially well investigated (review: Maru, 2001). The *Philadelphia translocation* is a chromosome translocation affecting the *c-abl* gene of chromosome 9 and the *bcr* gene of chromosome 22. The translocation leads to the formation of a hybrid gene composed of the *bcr* gene, which codes for a Ser/Thr-specific protein kinase, and the *c-abl* gene. Consequently, the two alternative fusion proteins $p210^{BCR\text{-}ABL}$ and $p180^{BCR\text{-}ABL}$ are created, which are characteristic of various leukemias.

During the translocation, a part of the *c-abl* gene is fused to the first exon of the *bcr* gene (Fig. 14.5). The $p180^{BCR\text{-}ABL}$ hybrid protein demonstrates increased tyrosine kinase activity, and it has a changed subcellular location in that it is predominantly found in the cytosol, whereas c-Abl normally exerts its function in the nucleus. Nuclear c-Abl has been shown to have an important role in damage-induced apoptosis. It is found in a repressed state bound to retinoblastoma protein pRb during G_o and G_1. DNA damage releases c-Abl from the repressed state, and the activated c-Abl induces cell death in a pathway involving p53 and the p53-related protein p73.

The Bcr/Abl hybrid is found exclusively in the cytosol, where it activates pathways normally under the control of receptor tyrosine kinases. Among the pathways activated by cytoplasmic Bcr/Abl are the MAPK, Jak/Stat, and PI3 kinase pathways.

Regulatory GTPases

Oncogenic activation of small regulatory GTPases has been documented many times for the example of the Ras proteins (see Section 9.2.3), and mutated Ras proteins have been detected in a large fraction of solid tumors. Studies in the transforming mechanisms of oncogenic Ras have revealed a complex cooperation of Ras activation with other signaling pathways (review: McCormick, 1999). Oncogenic mutants of Ras transmit prolonged and persistent signals to the downstream effectors, among which the Raf kinase, PI3-kinase, and RalGDS are the most important. In cellular model systems of Ras transformation, these Ras effector pathways have been shown to act synergis-

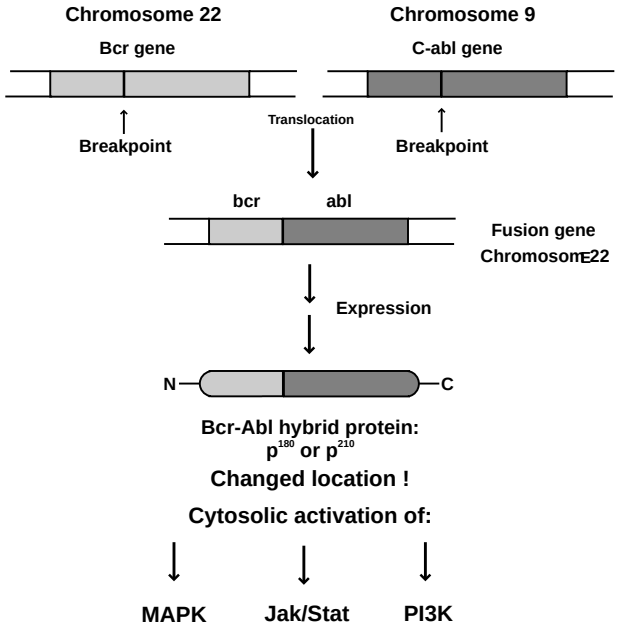


Fig. 14.5 Formation of a hybrid oncoprotein, illustrated by translocation of the Abl tyrosine kinase. The gene for the Ser-specific protein kinase BCR is fused with a part of the *c-abl* gene in the process of the Philadelphia chromosome translocation. Fusion genes are produced on chromosome 22, coding for various fusion proteins. The most important fusion proteins are the p^{180}- and p^{120}-BCR-Abl hybrid proteins, which have increased Tyr kinase activity and an altered subcellular location.

tically. In the normal situation, however, the pathways operate more independently and show precise temporal fluctuations. High levels of constitutively activated Ras are thought to engage effector proteins that are not activated by normal Ras. As an example, Ras plays a minor role in normal signaling to PI3-kinase from growth factor receptors such as the platelet-derived growth factor receptor, PDGF-R. However, oncogenic Ras is a strong activator of the PI3 kinase pathway. As outlined in Section 6.6.1, PI3 kinase activation creates a signal for increased cell survival.

Oncogenic activation of heterotrimeric G-proteins, in contrast, is seldom observed. In some tumors of endocrine organs (thyroid glands, pituitary gland), mutated $G_{s,\alpha}$ subunits occur which have a strongly reduced GTPase activity. Consequently, there is a constitutive activation of cAMP, which sets in motion uncontrolled cell division in the affected cell types. The mutations of the α-subunits affect the positions Arg201 and Gln227. Arg201 is at the site of ADP ribosylation by cholera toxin (see Section 5.5.2). The Gln227 is equivalent to Gln204 of the G_i subunits. It is directly involved in the GTPase reaction (see Section 5.5.4).

Cyclins

Oncogenic activation of cyclins is mostly observed for the D-type cyclins, which play a central role in the transition from G_o to G_1 and for G_1 progression. Increased levels of D-type cyclins and of CDK4/6 activity are frequently found in tumors because of gene translocation, overexpression, and amplification of the cyclin D genes. Furthermore, hypomethylation of the gene for cyclin D1 has been reported for tumors. Inappropriately elevated levels of D-type cyclins will increase the number of cells that leave G_0 phase and enter into G_1 phase. Furthermore, high levels of D-type cyclins make cells more independent of nutrient supply and will therefore add another growth advantage to cancer cells.

We know of several mechanisms that enhance cyclin D activity. Expression of cyclin D1 is, e. g., stimulated by the Ras/MAPK pathway and is also regulated via the APC pathway (see below, Section 14.9). Furthermore, cyclin D levels are controlled by ubiquitin-dependent proteolysis, by subcellular localization, and by phosphorylation (see Section 13.2.4).

Overexpression of cyclin D1 is also observed in some forms of bowel and breast cancer. The overexpression is mostly due to amplification of the gene for cyclin D1.

Cylin E is less frequently found mutated in cancers. This is probably because of the critical function that cyclin E has for entry into S-phase. Mutations of cyclin E will often destroy this central function, induce apoptosis, and not allow further cell cycle progression.

Transcription Factors

A large number of proto-oncogenes code for transcription factors required for progression of the cell cycle and/or for the differentiation of the cell. The best-known and most thoroughly investigated examples of oncogenic mutated transcription factors involve the *jun*, *fos*, and *myc* genes, and the genes for the T_3 receptor and the vitamin A acid receptor. Of these, aberrant activity of the Myc transcription factor has frequently been detected in tumor cells, and Myc is now considered to be one of the major oncogenes in real tumors. The Myc transcription factor controls a multitude of genes important for cell cycle regulation (Menssen and Hermeking, 2002). Among these are the genes for cyclin D1, cyclin B1, CDK4, BRCA1, ARF and others. Therefore, Myc is an important control element for G_1 progression, and it appears to couple DNA replication to processes preserving the integrity of the genome. Furthermore, Myc has been shown to function – most probably indirectly – also as a repressor of transcription of genes, e. g., the gene for the inhibitor $p15^{INK4b}$.

The Myc protein belongs to the family of basic helix-loop-helix proteins (see Section 1.2.1) and forms heterodimers via its leucine zipper motif with another helix-loop-helix motif transcription factor, the MAX protein. The formation of Myc-Max heterodimers is a prerequisite for binding to the cognate DNA element in the promotors of the Myc target genes and activation of these genes. The Max protein can heterodimerize with other helix-loop-helix proteins like the MAD protein or the Mxi-1 protein. The various heterodimers have different effects on transcription activity of the corresponding genes. The c-Myc-MAX dimer activates transcription, while the MAX-MAD or MAX-Mxi-1 dimers repress rather than activate transcription. In a normal cell, there is a balanced equilibrium between the different dimers.

In the well-studied example of Burkitt's lymphoma, Myc expression is upregulated in the course of a chromosome translocation (Fig. 14.6). Oncogenic activation of transcription factors due to chromosome translocation is often observed in leukemias. In Burkitt's lymphoma, translocation brings the *myc* gene into the vicinity of immunoglobulin genes. The translocation has the consequence of increasing expression of the *myc* gene in comparison to the normal situation. Furthermore, during the course of tumor formation, mutations also occur in the coding region of the *myc* gene. The Myc-Max equilibrium is disturbed by the chromosomal translocation and the associated overexpression of the Myc protein (see Fig.14.6) . It is assumed that there is an excessive transcription of target genes due to overexpression of the Myc protein, leading to tumor formation..

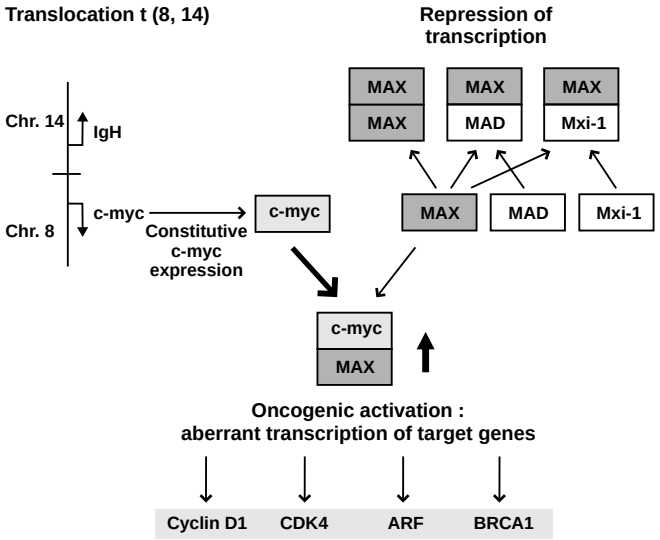


Fig. 14.6 The network of the c-Myc-Max transcription factors in Burkitt lymphoma. Burkitt lymphoma is characterized by chromosome translocations in which immunoglobulin gene sections (a gene for IgH in our example) are translocated into the region of the *c-myc* gene. As a consequence of the translocation t(8,14), constitutive expression of the *c-myc* gene occurs. The *c-myc* gene codes for the transcription factor c-Myc, which can form homodimers or may associate with the related transcription factors MAX, MAD and Mxi-1 to form heterodimers. Constitutive expression of c-Myc shifts the homo-heterodimer equilibrium towards the c-Myc-MAX heterodimers. Unprogrammed activation of target genes (genes for cyclin D1, BRCA1, CDK4, ARF) of the cMyc-MAX heterodimers then occurs.

14.4 Tumor Suppressor Genes: General Functions

Tumor suppressor genes are roughly defined as genes that have a negative, suppressing effect on tumor creation and thus help to prevent the formation of tumors. Mutations of tumor suppressor genes are often recessive. On the mutation of one allele, the remaining intact allele on the other chromosome continues to perform the growth-suppressing function. Only when both alleles are inactivated does the tumor-suppressing function cease to work. By this property, the inactivation of neighboring marker genes during tumorigenesis (loss of heterozygosity, LOH) has often helped to identify tumor-suppressor genes. However, tumor-suppressor genes are known for which mutation of only one allele promotes tumor formation. Halving the gene dose is apparently sufficient in these cases to lift tumor suppression. An example is the cell cycle inhibitor $p27^{KIP1}$ (Fero et al., 1998). Another mechanism of great importance for tumor formation is the epigenetic silencing of tumor suppressor genes by aberrant DNA methylation (see Section 13.1.3).

Based on the functions in tumor formation, the tumor suppressor proteins can be roughly divided into two classes (Kinzler and Vogelstein, 1997):

- **"Gatekeeper"**
  This type of tumor-suppressor protein directly prevents growth of tumors by inhibiting cell growth or promoting apoptosis. Each cell type has only a few gatekeepers, and inactivation of these genes leads directly to neoplastic growth. These genes normally function as gatekeepers to prevent uncoordinated growth. Important gatekeepers are the pRb protein, the p53 protein, and the cell cycle inhibitor $p16^{INK4a}$. Inactivation of these genes is rate determining for tumor formation, and both the maternal and the paternal copies of the gene must be inactivated for tumor formation.
- **"Caretaker"**
  Genes and proteins known as caretakers have an indirect influence on tumor formation. These are susceptibility genes that indirectly suppress tumor formation by maintaining the integrity of the genome. Abrogation of the caretaker function will enhance genetic instability and will allow the accumulation of mutations. An important class of caretakers includes repair proteins – components of DNA damage checkpoints. Inactivation of a caretaker gene of this class leads to a sharp increase in mutation rate and is therefore equivalent to constant exposure to mutagens.

Supression of tumor formation can be achieved by processes at various levels, including DNA repair, repair checkpoints, cell cycle regulation, apoptosis, and processes that are required in the later stages of tumorigenesis, e. g., blood vessel formation. Loss or deregulation of these functions by mutations or by epigenetic processes will favor tumor formation. Defects in repair genes or in repair checkpoints will allow the accumulation of further mutations and will expand the pool of variants on which selection will occur. Loss of growth-suppressing functions provides a growth advantage for the mutated cells and leads to deregulated proliferation. Inactivation of apoptotic responses will favor the survival of cells that would die in the normal situation.

Inactivation of tumor suppressor genes can have various consequences:

- **Loss of negative, suppressing signals in cell division**
  Mitogenic signal transduction pathways and the cell cycle machinery contain a range of negative regulation elements, which help to reduce or terminate a mitogenic signal (see Chapter 13). Inactivation of the suppressing function is equivalent to stimulation of cell division in many cases. Negative control of cell division activity is performed in particular by the repression of genes with products that are required for progression in the cell cycle. The retinoblastoma protein pRb and the CDK inhibitors are of special importance here.
- **Promotion of the accumulation of further mutations**
  A cell that has suffered DNA damage can bring about a halt in the cell cycle with the help of intrinsic control mechanisms (DNA damage checkpoints, see Section 13.8). The aim is to gain time for repair of the DNA damage and to prevent cells with the DNA damage progressing in the cell cycle. Failure of the regulatory functions that couple progression in the cell cycle to integrity of the DNA will favor the establishment of DNA damage as a mutation. The mechanism by which coupling is achieved between DNA damage and progression of the cell cycle is only just starting to be understood. The tumor-suppressor protein p53 has a central function at DNA damage checkpoints (see Section 14.8.6).
- **Failure of apoptotic signals**
  Apoptosis or programmed cell death is a program that brings about the death of the cell in a targeted manner in the presence of DNA damage (see Chapter 15). Apoptosis is a protection against the formation of tumor cells. If a cell is affected by DNA damage, apoptosis can help to initiate cell death before further mutations accumulate that would favor transition to the tumor state. Many tumor-suppressor proteins are involved in the induction of apoptosis, and failure of this function will allow the survival of cells that would normally be eiliminated by the apoptotoc program.

A selection of well-characterized tumor suppressor proteins is given in Table 14.2. Most of the tumor suppressor genes can be categorized according to the functions listed above. However, a number of tumor suppressor genes are known with no direct relationship to the regulation of the cell cycle, repair, or apoptosis. Some of the tumor suppressor genes in Table 14.2 are involved in the organization of the cytoskeleton or in cell-cell interactions and appear to be relevant in later stages of tumor formation where tumor cells invade foreign tissues and form organ-like structures.

14.5
DNA Repair, DNA Integrity and Tumor Suppression

Each cell has a range of protection mechanisms to avoid DNA damage and to ensure the integrity and stability of the genome. The following examples of caretakers are associated with tumor creation and tumor progression.

Tab. 14.2 Characteristics of some tumor suppressor proteins.

Gene, protein	Function
p53	DNA repair, apoptosis
pRb	Cell cycle control
NF1, neurofibromin	, GAP in Ras signaling
BRCA1, BRCA2	DNA repair, e. g. of double strand breaks
Wt-1	Transcription factor with Zn-binding motif
PTEN	Phosphatidylinositol phosphate phosphatase , blocking of PI3-kinase signaling
CDH1, E-Cadherin	cell adhesion
APC	Binds to ß-catenin; Wnt-signaling
MSH2,	Repair of DNA mismatches

Mismatch Repair, hMSH2

For inherited forms of a certain form of bowel cancer (*hereditary nonpolyposis cancer*, HNPPC), it has been observed that there is an error in the function of the repair system for DNA mismatches. Patients with HNPPC have inherited a defect in the hMSH2 gene in their germ cells, and their tumor cells have a further mutation in the hMSH2 gene. The hMSH2 gene is a homolog of the MutS gene in *E. coli*, and its gene product is involved in the repair of DNA mismatches. The defect in the mismatch repair is responsible for a type of genetic instability most readily observed at the level of microsatellite DNA (see also Section 14.1.5). This DNA harbors a large number of repeat sequences, and defects in mismatch repair lead to easily detectable changes in the number and sequence of the repeats.

The BRCA Genes

Two genes with tumor-suppressing function, BRCA1 and BRCA2, are known that mediate a hereditary susceptibility to breast cancer. Inactivation of BRCA1 function can occur by mutation of the gene as well as by epigenetic silencing due to promotor hypermethylation. The latter is frequently observed in sporadic breast cancers.

BRCA1. The product of the BRCA1 gene is a large protein of 1883 amino acids which harbors functions related mostly to DNA repair. The following processes have been linked to the BRCA1 protein:

- *DNA recombination*
  BRCA1 interacts specifically with the Rad51 protein which catalyzes strand exchange during recombination processes. By this property, BRCA1 is assumed to participate in DNA double-strand break repair. Furthermore, BRCA1 shows a DNA-binding activity specific for branched DNAs.
- *DNA damage checkpoint*
  BRCA1 is essential for activating the protein kinase Chk1 that regulates DNA damage-induced G2/M arrest (review: Yarden et al., 2002).
- *Transcriptional regulation*
  BRCA1 contains a transactivating domain and appears to activate the transcription of essential growth-controlling genes like the genes for the transcription factor c-

myc, for cyclin D1, for the Stat1 protein, and for the protein kinase Jak1. In addition, BRCA1 can repress the activation function of the estrogen receptor (Zheng et al., 2001).

- *Ubiquitin-mediated proteolysis*
  A RING finger motif (see Section 2.6.1) and ubiquitin ligase activity have been identified in BRCA1. The purpose of the E3 activity of BRCA1 is unknown.

BRCA2. This protein also interacts with Rad51, and it is assumed to participate in the biological response to DNA damage, mainly double-strand break repair of DNA.

Defects in the Metaphase-Anaphase Checkpoint

The metaphase-anaphase checkpoint controls the formation of the spindle apparatus and the correct alignment of chromosomes, and may initiate metaphase arrest (see Chapter 13). On failure of the control system, the occurrence of abnormal chromosomes is favored. In various cancer cells, chromosomal instability is associated with loss of function of the BUB1 gene, which is part of the metaphase-anaphase checkpoint (review: Ru and Chen, 2001).

14.6 The Retinoblastoma Protein pRb in Cancer

The retinoblastoma gene was the first tumor suppressor gene to be identified and characterized in man. Human genetic investigations of patients suffering from the rare retinoblastoma eye tumor showed a defect in a gene sequence known as the retinoblastoma gene. In the inherited form of this tumor, which appears in childhood, a defect is inherited in the retinoblastoma gene via the germ line. Inactivation of the second allele of the retinoblastoma gene by mutation or deletion leads to complete failure of the function of the gene and thus to tumor formation. Somatic mutations of the *pRB* gene are found in more frequently occurring tumors, including osteosarcomas, lung carcinomas, and bladder carcinomas.

The product of the retinoblastoma gene, the retinoblastoma protein pRb, functions as a switch for the conversion of mitogenic and antimitogenic signals at the transcription level in the cell cycle (see Section 14.3.2). The switch function is performed with the help of the phosphorylation status of pRb (see Fig. 13.16). In the hypophosphorylated state, and in cooperation with the transcription factor E2F, pRb represses the expression of genes with products needed for progression through the cell cycle. In the hyperphosphorylated state, in contrast, the pRb protein has an activating effect on these genes. The phosphorylation status of pRB (and its relatives, see Section 13.4.2) is controlled by a network of interactions between D-type cyclins, CDK 4/6, cyclin E/CDK2, and the CDK inhibitors, mainly of the INK4 family.

It is now increasingly clear that a defect in the pRB pathway has an effect on both cell proliferation and apoptosis. Cell proliferation will increase when a mutated, inactive form of pRb no longer keeps the E2F-controlled genes repressed. In this case, uncon-

trolled activation of E2F takes place, and an increased supply of gene products necessary for S-phase progression provides a growth advantage for the mutated cell.

A function equally important for tumorigenesis is the ability of pRb to protect differentiating cells, which contain high levels of pRb, from apoptosis (Fig 14.7). The link from the pRb pathway to apoptosis appears to be provided by E2F1, which controls not only genes required for S-phase progression but also genes involved in the regulation of apoptosis. Among the latter are the genes for the pro-apoptotic protein Apaf1 (see Section 15.5) and for the tumor suppressor ARF (see Section 14.7), which is part of the network that regulates p53 function. Loss of pRb function – in the absence of other mutations – has been shown by knockout studies in mice to result in increased apoptosis and in growth arrest. If at the same time other key regulators of apoptosis, e. g.,

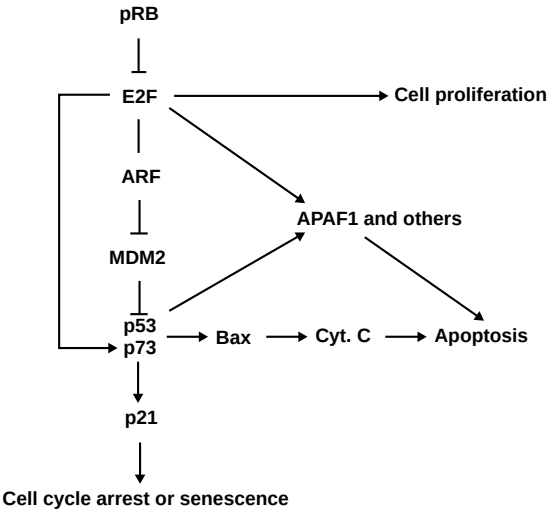


Fig. 14.7 Model of pRb function in control of cell proliferation and apoptosis.
pRB controls controls proliferation and apoptosis through active repression of E2F-dependent promotors. In the underphosphorylated form, pRb has both an anti-proliferative and anti-apoptotic effect. It inhibits transcription of E2F-controlled genes, among which are genes required for S-phase progression and hence for cell proliferation. E2F-controlled genes also include pro-apaoptotic genes such as the gene for Apaf1 and for p73. The latter can induce, in addition to p53, the gene for the proapoptotic protein bax which stimulates cytochrome c release from mitochondria and thereby initiates apoptosis. E2F also keeps levels of p53 low, an effect mediated by the ARF-MDM2-p53 network (see Section 14.8.6). If the control function of pRb is lost e. g. due to pRb mutations, E2F will inappropriately activate the pro-apoptotic genes and growth arrest and/or apoptosis will result. When at the same time pRb function is lost and key players regulating apoptosis and /or cell cyle arrest (e. g. ARF, p53, Apaf1) are mutated or overexpressed (MDM2), cells will be directed to hyperproliferation.

the p53 protein or ARF, are mutated and inactivated, apoptosis is reduced and cells are driven to hyperproliferation. In accordance with this model, nearly all cancers contain mutations both in the pRb pathway and in the pathways that link proliferation with apoptosis. This example illustrates the intense networking that exists between the various growth- and cell death-controlling pathways, and it underlines the cooperativity of mutations in key regulators of cell survival and cell death.

The regulatory network, of which the pRb protein is a part, can be changed by various mechanisms in the sense of stimulation of proliferation (Fig. 14.8). First, an error in the function of proteins involved in the phosphorylation of pRb may initiate a dominating activation of the pRb function. Second, the pRb protein itself may be activated in an unprogrammed manner by mutation or by binding of viral proteins.

Unprogrammed Activation of Cyclins and CDKs

Inactivation of the repressing function of pRb can be achieved by an uncontrolled increase of D-type cyclins and CDK2/4, leading to inappropriate phosphorylation of pRb and subsequent crossing of the restriction point. Overexpression of cyclin

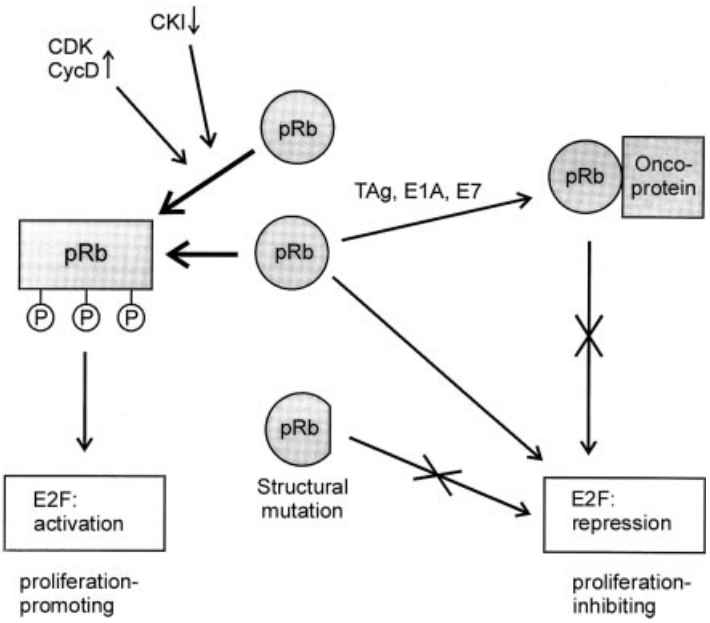


Fig. 14.8 Errors in regulation of the tumor suppressor protein pRb. The figure shows a simplified version of well characterized mechanisms by which errors in regulations of the pRb function can occur. The proliferation-promoting activity of pRb is amplified if unprogrammed high activity of the CDK4-cyclin D1 complex is present or the inhibitory activity of CKI is absent. Lack of proliferation-inhibiting activity of the pRb protein can be due to complexation of the pRb protein with viral oncoproteins or due to structural mutation of the pRb protein. The transcription factor E2F no longer has a repressing activity on transcription of cell cycle genes in this situation (see 14.3) and the division-inhibiting activity of the pRb protein is absent.

D1 or CDK4 has been found in various tumors of the bladder, breast, pancreas, and lungs.

Inactivation of Inhibitors of CDKs (CKIs)

CKIs inhibit the phosphorylation of pRb and bring about a halt in the G_1 phase. Of the various CKIs, mutations in the inhibitor $p27^{KIP1}$ and, in particular, the $p16^{INK4a}$ protein (see Section 13.2.6) are associated with tumor formation. In many tumors, such as lung carcinomas, inactivation of the gene for the $p16^{INK4a}$ inhibitor has been observed, based on mutations or an aberrant C-methylation. Low levels of the $p16^{INK4a}$ inhibitor promote the transition of cells from the quiescent to the proliferating state and enhance cell proliferation.

Binding of Viral Oncoproteins

DNA viruses that can trigger tumors are found in the classes of the polyoma viruses, the adenoma viruses, and the papilloma viruses. The polyoma viruses, with the SV40 virus as a well-studied representative, the adenoma virus, and the human papilloma virus (HPV) are associated with the formation of tumors in humans and have genes coding for proteins with the properties of oncoproteins. The oncoproteins of all three viruses interfere with the pRb function by lifting its inhibition of transcription factor E2F. It is assumed that the tumor-promoting activity of the proteins is due, in particular, to this property. Interestingly, these proteins also bind to and inactivate the tumor suppressor protein p53. As outlined above, mutant p53 can cooperate with mutant pRb in driving tumor progression.

The oncoproteins are the T antigen (TAg) of the SV40 virus, the E1A protein of the adenoma virus, and the E7 protein of HPV. The three proteins have in common the ability to bind to the hypophosphorylated form of pRb. In all three cases, binding takes place in the same region of pRb referred to as the "pocket". The transcription factor E2F also binds in the region of the "pocket" of pRb protein. Binding of the viral proteins to pRb is thought to release E2F from repression, allowing unprogrammed transcription of the E2F target genes.

Genetic Inactivation of pRb

Genetic inactivation of pRb is observed in many tumors. The gene defect may affect the promotor region of the pRb gene, leading to reduced pRb expression, or it may affect the structure of pRb, for example, by a mutation of the binding site for E2F. The mutations observed in tumors are generally extensive structural changes in the pRb gene.

14.7 The $p16^{INK4a}$ Gene Locus and ARF

Investigations into tumor cells indicate that proteins coded by the $p16^{INK}$ gene locus function as tumor suppressors and are of great importance for tumor development (review: Sherr, 2001). The gene locus for $p16^{INK4a}$ codes for two proteins, namely

the $p16^{INK4a}$ inhibitor and the ARF protein (ARF = alternative reading frame: $p19^{ARF}$ in mice; $p14^{ARF}$ in man). Both proteins have a growth-inhibiting function, though with different points of attack. Whilst the $p16^{INK4a}$ protein inhibits the cyclin D-CDK complex and brings about a halt in the cell cycle via the pRb protein, the ARF protein attacks the function of the p53 protein by specifically interacting with the MDM2 protein and interfering with its binding to the p53 protein (see Section 14.8.6, Fig. 14.14). The ARF protein is not homologous to the $p16^{INK4a}$ protein, although both originate from the same gene locus. It arises by alternative splicing and by use of a different reading frame.

An important link exists between the pRb/E2F pathway, since ARF is one of the transcriptional targets of the E2F transcription factors. Furthermore, ARF is under transcriptional control by the transcription factor c-Myc, providing a link from c-Myc to the function of p53.

14.8 The Tumor Suppressor Protein p53

The most frequently observed genetic changes in human tumors affect the gene for a nuclear phosphoprotein of 393 amino acids, which is known as the p53 protein, after its molecular weight.

Mutations of the p53 gene are observed in over 50 % of all human tumors. Defects in the p53 gene in the germ line lead to a hereditary tendency to develop various tumors, especially of the connective tissue. In affected families, several members of the family may develop tumors during childhood. The disease is known as Li Fraumeni Syndrome after its discoverer.

The close correlation between tumor formation and mutations of the p53 gene indicates that the p53 protein has a central function in tumor pathogenesis. Thanks to intensive investigations of the structure and functions of the gene and its coded proteins, it is clear that the p53 protein is an important component of a regulatory network in which cell cycle control, integrity of the DNA, and programmed cell death play a central role.

The following central functions may be assigned to the p53 protein (review: Guimaraes and Hainaut, 2002):

- The p53 protein functions as a sequence-specific transcription activator. As such, it is able to bind specifically to DNA elements and to activate the transcription of downstream genes with functions in cell cycle control, apoptosis, angiogenesis, and stress response.
- The p53 protein is part of a control mechanism that couples progress of the cell cycle to the integrity of the DNA. In the presence of DNA damage, the p53 protein can bring about a halt in the cell cycle at the G1→S transition and at other points of the cell cycle.
- The p53 protein is involved in the initiation of programmed cell death.

14.8.1
Structure and Biochemical Properties of the p53 Protein

The p53 protein can be seen as a multi-talent amongst the regulatory proteins. As shown in Fig. 14.9, several domains can be identified in the p53 protein, and defined biochemical functions can be assigned to these (Giaccia and Kastan, 1998).

1. **Transactivation domain**
   In the N-terminal region of p53, there is a transactivation domain which p53 uses to make contact with the transcription apparatus. Different protein-binding sites have been identified in this region. These include binding sites for components of the TFIID complex and for coactivators such as the histone acetylases CBP/p300 or PCAF (see Section 1.4.7.1).

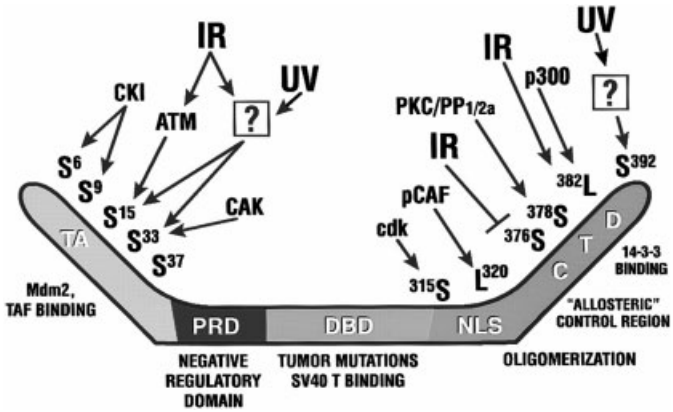


Fig. 14.9 Posttranslational modifications and localization of biochemical functions in p53 protein. The sequence sections and the sites of posttranslational modifications of the p53 protein, to which a biochemical function has been assigned, are shown .
The major domains of p53 characterized as transcriptional activation (TA), proline-rich domain (PRD), DNA-binding domain (DBD), nuclear localization signal (NLS), and carboxy-terminal domain (CTD).
The TA is also the region where Mdm2 and TAF [TATA-binding protein (TBP)-associated factor] bind. PRD may be a negative regulatory domain. The DBD is the site of the vast majority of tumor mutations and is the area where SV40 large T antigen binds. The CTD is considered the region of major allosteric regulation of p53 function and contains the sequences necessary for dimerization and tetramerization. Casein kinase I (CKI) phosphorylates serines 6 and 9 in vitro. ATM and DNA-PK phosphorylate serine 15 in vitro and ionizing radiation (IR) activates ATM kinase activity. CDK activating kinase (CAK) phosphorylates serine 33 in vitro and IR induces serine 33 phosphorylation. Cdk phosphorylates serine 315 in vitro. PKC phosphorylates and phosphatases 1 and 2a dephosphorylate serine 378 in vitro. Histone acetylases pCAF and p300 acetylate in vitro lysine 320 and lysine 382, respectively. IR induces lysine 382 acetylation and serine 376 dephosphorylation. UV induces phosphorylation of serines 15, 33, and 392. After Giaccia and Kastan, 1998.

2. **Sequence-specific DNA-binding domain**
   A core domain of 100–300 amino acids includes the binding site for the corresponding DNA element and binding sites for viral oncoproteins such as the large T antigen (TAg) of SV40 virus.
3. **Basic C-terminal domain**
   The C terminus of p53 contains a basic domain where DNA can be bound in a nonspecific way. Furthermore, the C-terminal domain contains several sites for post-translational modifications, including Ser-phosphorylation, Lys-acetylation, ubiquitination, and Lys-sumoylation. Futhermore, sequence signals for nuclear localization, sequence sections for tetramerization, and binding sites for transcription factors are found in the C-terminal part. Overall, the C-terminal domain has an important function for the regulation of p53. There is experimental evidence that specific DNA binding of the core domain is controlled by phosphorylation of the C-terminal domain.

14.8.2
Sequence-Specific DNA Binding of p53

Central to the function of the p53 protein is its ability, as a transcription activator, to specifically bind to corresponding *cis* elements in the promotor region of various genes and to activate their transcription. The importance of sequence-specific DNA binding for the tumor-suppressing function of p53 became clear when the crystal structure of the complex of p53 protein and a corresponding DNA element was resolved and this structure was compared with the spectrum of known mutations of p53 protein occurring in human tumors (Cho et al., 1994).

The p53 protein is a tetrameric protein that binds to DNA elements with the consensus sequence 5'-RRRC(A/T)(T/A)GYYY-3' (R = purine, Y = pyrimidine). The structure of the complex of the central DNA-binding domain of p53 with an oligonucleotide that carries one half-site of the p53 recognition sequence is shown in Fig. 14.10.

The p53 protein contacts the recognition sequence by two means. A loop-helix-loop motif is placed in the large groove of the DNA and makes contact with the bases. In addition, another loop (L3) forms a contact via an Arg side chain (R248) to the minor groove of the DNA.

The large number of known sequences of the p53 gene from tumor patients was particularly valuable for interpretation of the crystal structure, since a spectrum could be assembled for p53 mutation in association with tumor formation. The mutation spectrum shown in Fig. 14.11 shows "hotspots", positions at which p53 mutations are seen particularly frequently in tumor patients. These hotspots cluster in the core domain of p53 responsible for sequence-specific DNA binding. Remarkably, few oncogenic mutations are found outside of the DNA-binding domain.

Comparison of the *mutation spectrum* with the structure of the p53-DNA complex indicates that the positions of frequent mutations coincide with the conserved structural elements of the DNA-binding domains. It is particularly noticeable that the most frequently mutated position in tumors, namely Arg248, is also the position at which the p53 protein forms a specific contact to recognition sequences.

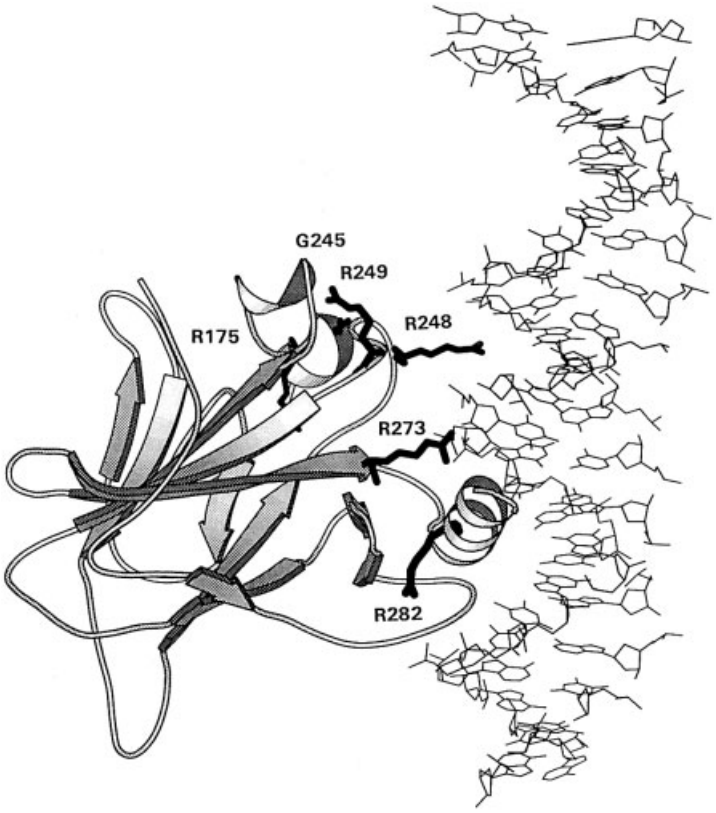


Fig. 14.10 DNA binding domain of the tumor suppressor protein p53 in complex with DNA. Crystal structure of the core domain of p53 (amino acids 102 - 292) in complex with a double-stranded DNA that contains a specific binding site for p53 (Cho et al., 1994). The amino acid positions are highlighted at which frequent oncogenic mutations are observed (see Fig. 14.11).

The naturally occurring mutations in the DNA-binding domain can be divided into two classes, as described below.

The first class includes mutations of amino acids that enter into direct contact with the DNA element. These include mutations at positions 248 and 175. Both positions show a high mutation frequency in naturally occurring tumors.

The other class includes mutations that are assumed to change the overall structure of the DNA-binding domain of p53 so that a specific interaction is no longer possible with the DNA.

Overall, comparison of the structure data with the natural mutation spectrum of p53 shows that specific DNA binding is closely associated with the function as a tumor suppressor. A disruption of specific DNA binding apparently has serious effects on the tumor-suppressing function of p53.

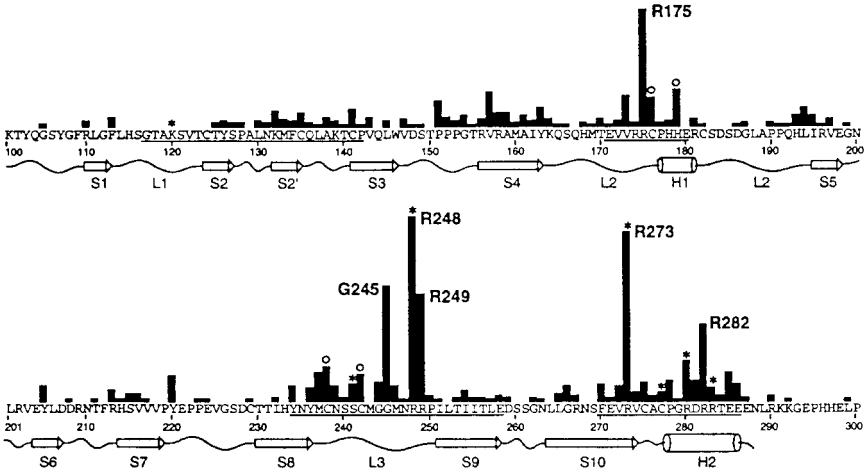


Fig. 14.11 Mutation spectrum of the p53 protein in tumors. The linear structure is shown of p53 and the frequency of mutations found in tumors. The black bars indicate the approximate position and the relative frequency of the p53 mutations. The frequency of mutations in the region of the DNA binding domain is of note. The sites of the most frequent mutations coincide with positions of the p53 protein that are directly involved in interactions with the DNA sequence (see Fig. 14.9).

14.8.3 Genes Regulated by p53

The p53 protein functions as a specific transcription activator, but it can also bring about a repression of distinct genes. Table 14.3 gives an overview of genes regulated in response to p53 activation.

1. **Activation of specific genes**
   The genes activated by p53 can be grouped by their biological function:
   - **Coupling of DNA damage to cell growth arrest**
     The products of these genes have a suppressing effect on cell cycle progression in case of DNA damage, and their levels are increased upon p53 activation. Loss of p53 function will lower the threshold for cell cycle arrest and will allow cell cycle progression in the presence of damaged DNA.
   - **$p21^{CIP1}$**
     The most prominent example is the CDK inhibitor $p21^{CIP1}$. Activation of p53 leads to increased formation of the p21 inhibitor, which brings about a halt in the cell cycle at G1/S and and G2/M.
   - **GADD45**
     GADD45 functions as an inhibitor of PCNA during DNA replication and brings about a growth arrest during repair processes.
   - **14-3-3 σ**
     This protein has an important function in the G_2/M DNA damage checkpoint. It serves to sequester the mitotic initiation complex, CDC2-cyclin B1, in the cyto-

Tab. 14.3 Examples of genes and proteins regulated in response to p53 activation.

Target genes	Biological function
$p21^{Cip1}$ GADD45 14-3-3 sigma Cyclin A	Cell growth arrest
BAX Bcl-2 AIP1 CD95(Fas) NOXA PUMA	Apoptosis
XPB and XPD Rad51 TFIIH p62 subunit	DNA repair
TSP1, thrombospodin1 gene	Angiogenesis
COX2, cylooxygenase 2 NOS1, nitric oxide synthase 2 GPX, glutathion peroxidae	Stress responses
MDM2	Retrocontrol, G_2/M

plasm after DNA damage. Levels of 14-3-3 σ have been found to be low in many tumors, including breast tumors.

- **Apoptosis**
  A growing list of genes involved in the induction and course of apoptosis have been identified as transcriptional targets of p53. Of note are the genes for the proapoptotic protein Bax (see Section 15.4), which is activated by p53, and the gene for the antiapoptotic protein Bcl-2 (see Section 15.4), which is repressed by p53. The latter effect indicates that p53 can promote apoptosis by blocking survival signals mediated by Bcl-2. Induction of the death receptor ligand Fas/Apo1 (see Section 15.6) has also been reported, suggesting that p53 may induce caspase activation through death receptor signalling.
- **Stress responses**
  Various stresses (oxidative stress, hypoxia, ribonucleotide depletion) lead to p53-mediated activation of genes involved in stress responses. Among these are many genes that can generate or respond to oxidative stress.
- **Angiogenesis**
  The thrombospodin-1 gene codes for a protein that inhibits new formation of blood vessels (angiogenesis). The p53 protein activates expression of the thrombospodin-1 gene and can thus suppress angiogenesis. If the regulating activity of the p53 protein is inactivated, a situation is created which facilitates new formation of blood vessels, since the inhibitor of angiogenesis is missing. It

is assumed that this situation promotes tumor progression, especially in the late phase of tumor formation.

- **MDM2**
  The *mdm2* gene is activated by p53. Its encoded protein is part of a negative feedback loop most important for regulation of p53 activity (see Section 14.8.6).

2. **Transcription repression**
   The p53 protein has – in addition to the activation of specific genes – also a general repressing influence on transcription. The repression is observed for various cellular and viral genes that have no p53-binding site. Examples of genes repressed by p53 are the genes for transcription activators *c-jun* and *c-fos*, the cytokine IL-6, the retinoblastoma protein pRB (see Chapter 14), and the *bcl-2* gene (see Chapter 15). It is assumed that the general repressing activity of p53 is based on interference in the formation of the preinitiation complex of transcription and that this is associated with the binding of p53 to the TATA box protein TBP and the TFIID complex.

The large list of effectors of p53 activation shows that it controls a variety of pathways involved mostly in growth control and cell survival. The exact combination of effectors activated by p53 at the transcriptional level may differ from cell to cell depending upon the nature and intensity of the inducing signal and the cell type. Furthermore, the activation of other pathways contributing to the control of cell proliferation should also be considered.

14.8.4
Activation, Regulation and Modulation of the Function of p53

Activation and General Control Functions of p53

The p53 protein is present in almost all tissues in low concentrations and in a repressed state. Several types of stresses can lead to its accumulation and de-repression by post-translational modifications. These stresses can be broadly divided into three classes (Fig. 14.12);):

- **Genotoxic stress**
  DNA damage by UV, X-rays, carcinogens, cytotoxic drugs, etc.
- **Oncogenic stress**
  Aberrant activation of growth factor-signaling cascades
- **Non-genotoxic stress**
  Metabolic stress like hypoxia, depletion of ribonucleotides.

Activation of p53 serves to protect the cells against the potentially disastrous consequences of these stresses and occurs as a normal mechanism of defence against neoplastic transformation of the cell. Three major biological processes are influenced by p53 activation, namely cell proliferation and differentiation, DNA repair, and programmed cell death (apoptosis).

The activity and function of p53 are controlled by various post-translational modifications including ubiquitination, phosphorylation, acetylation, and sumoylation.

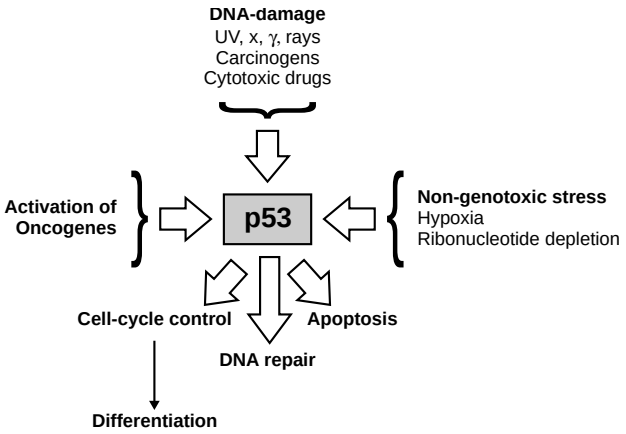


Fig. 14.12 Overview of the p53 pathway. The three main classes of signals activating p53 are shown After activation, p53 can exert a range of biological effects shown in the bottom part of the figure.

These modifications influence stability, subcellular localization, DNA-binding activity, and binding to effector proteins of p53 (see below):

- **Stability of p53 regulated by MDM**
  Activation of p53 in response to genotoxic stress is almost always accompanied by stabilization of p53. In the absence of stress, p53 levels are low because of degradation by the ubiquitin-proteasome pathway. Central to this degradation is the MDM2 protein, which specifically binds to an N-terminal region of p53, mediating its degradation in the ubiquitin pathway. In this process, MDM2 functions as an E3 ubiquitin ligase for p53. The interaction between p53 and MDM2 can be impaired by phosphorylation of p53 within the MDM2-binding region.
- **Subcellular localization of p53**
  Since one of the key functions of p53 is the regulation of transcription, localization of p53 to the nucleus plays an important role in the p53 response. p53 carries several nuclear localization signals as well as two nuclear export signals. Nuclear export appears to be an important control element of p53 activity. Binding of MDM2 enhances nuclear export and thereby diminishes the amount of p53 available for transcriptional activation, suggesting that nuclear export may be necessary for its ubiquitin-dependent degradation. Stress-induced phosphorylation of p53 at the N-terminus can inhibit MDM2 binding, which prevents ubiquitination and restricts nuclear export.
- **Sequence-specific DNA-binding activity of p53**
  About 18 phosphorylation sites have been identified on the p53 protein. These phosphorylations influence the interaction of p53 with transcriptional coactivators as well as its DNA-binding activity. Phosphorylation at the C-terminus has been shown to regulate the site-specific DNA binding located in the core of p53. The mechanism underlying this influence, however, remains to be clarified. Phosphor-

ylations at the N-terminus have also been shown to affect the transcriptional activity of p53. Binding between p53 and histone acetyl transferases like p300 or PCAF is enhanced by phosphorylation at specific sites of the N-terminus, without influencing MDM2 binding.

Acetylation of p53 at distinct Lys-residues of the C-terminus is observed concomitant with transcriptional activation. This modification has been reported to enhance the sequence-specific DNA binding. It remains to be shown whether this acetylation is a side effect of recruitment of histone acetylases or whether it serves directly to modulate sequence-specific DNA binding of p53.

- **Protein-protein interactions and p53 activation**
  In the course of p53 activation, numerous specific protein-protein interactions are engaged to perform the growth-controlling and pro-apoptotic functions of p53. Of these, the interaction with the MDM2 protein stands out, as discussed above. Other important protein-protein interactions include the binding of p53 to components of the basal transcription apparatus, e. g., the TFIIH p62 subunit, interaction with proteins of nucleotide excision repair (XPB, XPD), and the interaction with transcriptional coactivators.

  A special role is assigned to viral oncoproteins. The oncoprotein of SV40, TAg (see Section 14.4.4) binds to the p53 protein and can inactivate the p53 function by interfering with its transcriptional control activity.

14.8.5 Overview of p53 Regulation

The conditions that lead to activation of p53 can be broadly categorized into genotoxic stress, oncogenic stress and non-genotoxic stress. These stresses have been shown to initiate the typical p53 response, which can cause either apoptosis, growth arrest, altered DNA repair, or altered differentiation. Which of these responses dominates will depend on the cell type and the type and duration of the stress. It is becoming increasingly clear that the p53 function is embedded in a finely tuned regulatory network that uses various signaling pathways for the reception of activating signals and directs the p53 response to different downstream effector pathways. In this process, post-translational modifications of p53 are the major tools for the regulation of p53 function, and these modifications are directed and controlled by a multitude of input signals. Broadly, pathways leading to p53 activation and response can be categorized according to the initiating stresses:

- **Activation of p53 by genotoxic stress**
  Genotoxic stress, as manifested by, e. g., the formation of DNA adducts and DNA strand breaks, activates p53 mostly via DNA damage checkpoints. In this process, the cell cycle is arrested and time is gained for repair of the damaged DNA, or alternatively, the cell is driven into apoptosis. It is the main purpose of this control to prevent replication of damaged DNA, which is a potentially mutagenic process. The links between DNA damage and p53 activation are summarized schematically in Fig. 14.13.

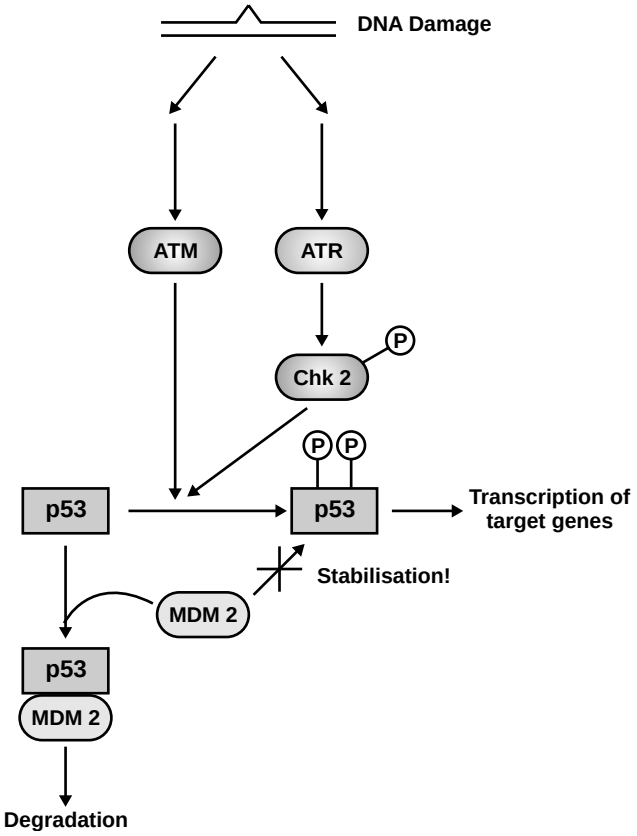


Fig. 14.13 Activation of p53 by DNA damage. DNA damage provides a signal for activation of the protein kinases ATR and ATM which phosphorylate p53 either directly of via another protein kinase, named Chk2. In the unphosphorylated form, p53 is instable due to binding of MDM2 which has ubiquitin ligase activity and induces ubiquitinylation and destruction of p53. The damage-induced phosphorylation of p53 prevents MDM2 binding and degradation leading to p53 accumulation and transcription of p53 target genes.

Upon DNA damage, a cascade of protein kinases is activated that catalyzes the phosphorylation of p53 on specific Ser residues, which, in turn, influence the interaction of p53 with the MDM2 protein. Specifically, p53 is phosphorylated on Ser 15 by the ATM and ATR kinases and on Ser20 by the Chk2 protein. All three kinases are part of DNA damage checkpoint pathways (see Section 13.8), and this phosphorylation inhibits binding of the MDM2 protein, leading to reduced degradation and higher steady-state concentrations of p53. The DNA damage-activated kinases enforce activation of p53, furthermore, by phosphorylating the MDM2 protein, which leads to enhanced ubiquitin-dependent degradation of MDM2.

It should be pointed out that proteins related to p53 and MDM2 have been identified that are also part of the MDM2-p53 regulatory module. Two proteins, named p63 and

p73, have been found to be "cousins" of p53. Both proteins are activated by similar stresses as p53 and elicit some, but not all, biological responses of activated p53. Furthermore, a relative of the MDM2 protein exists, which is named MdmX. The MdmX protein has been shown to modulate the ability of MDM2 to promote the degradation of p53.

- **Activation of p53 by oncogenic stress**
  The aberrant activation of proto-oncogenes creates a stress situation which is also termed oncogenic stress. Normal cells are protected against the detrimental consequences of oncogenic stress by linking oncogenic pathways to the pro-apoptotic action of the p53 pathway. Aberrant activation of proto-oncogenes to oncogenes results in activation of the p53 pathway and allows for the destruction of the cell by apoptosis. The main target of regulatory inputs during oncogenic stress is the MDM2 protein, which is downregulated under excessive survival signals, allowing for the accumulation and activation of p53 and subsequently the initiation of apoptosis.
  In this pathway, the ARF protein has a key function (Fig 14.14). The ARF protein binds to MDM2 and thereby reduces p53 degradation. Various signals induce an increase in ARF protein, including transcriptional activation via the pRb/E2F pathway, the Ras/MAPK pathway, and the myc transcription factor. As already pointed out in earlier chapters, these pathways receive and transmit a multitude of survival signals. By indirectly stabilizing p53, ARF plays a key role in eliminating cells that develop proliferative abnormalities, thereby protecting the organism from cancer development.
  In cells like stem cells with ongoing proliferation, however, the activation of the p53 response must be dampened to allow normal growth and development while retaining the capacity for induction of the response to stress associated with oncogenesis. One mechanism by which the p53 response is supressed in normal cells appears to use phosphorylation of MDM2 by the Akt kinase. The Akt kinase is activated by growth factor receptor pathways and transmits survival signals. Akt-mediated phosphorylation of MDM2 has been shown to promote nuclear localization of MDM2, increasing nuclear levels of MDM2 and promoting inhibition of the p53 response. Although an increasing number of key players of the regulatory network with the p53-MDM2 module at center stage have been identified, the quantitative aspects of this network and the cooperation of its components remain to be elucidated.
- **Activation of p53 by non-genotoxic stress**
  Non-genotoxic stresses like ribonucleotide depletion or under- or oversupply with oxygen can activate the p53 response without participation of the DNA damage checkpoints. The link between these stresses and p53 appears to be provided by a subspecies of the Jun-N-terminal kinase (JNK2), which is activated by various stresses and enhances p53 stability by phosphorylation on Thr81. Overall, however, this pathway is only poorly characterized.
- **Binding of viral oncoproteins**
  The oncoprotein of the SV40 virus, Tag (see Section 14.6), binds to the p53 protein and can inactivate the p53 function in a similar way to that assumed for inactivation of the pRb protein.

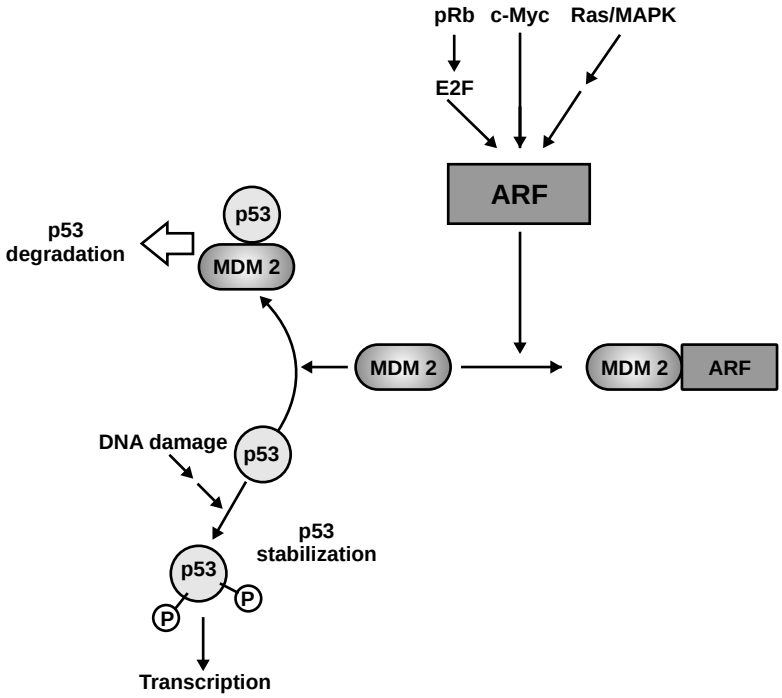


Fig. 14.14 Regulation of p53 activity by $p19^{ARF}$ and MDM2. A major pathway for downregulating p53 levels uses the MDM2 protein which is a ubiquitin ligase. MDM2 induces ubiquitinylation and degradation of p53. The downregulating function of MDM2 is mainly controlled by the tumor suppressor protein $p19^{ARF}$ which binds to MDM2 and thereby prevents p53-MDM2 binding and p53 destabilization. $p19^{ARF}$ transcription can be activated by the major mitogenic and proliferative pathways of the cell, namely the pRb/E2F pathway, the Ras/MAPK pathway and by the transcription factor Myc. By this mechanism, proliferative and mitogenic signals keep p53 levels low.

14.8.6 The MDM2-p53 Network and Cancer

A functional inactivation of the MDM2-p53 network (reviews: Alarcon-Vargas and Ronai, 2002; Michael and Oren, 2002, see Fig. 14.15) is observed in nearly all cancers. Inactivating mutations have been found to affect both the proliferative and the antiproliferative functions of the network, leading to increased cell division activity and to diminished apoptosis.

Activation of the proliferation activity by conversion of proto-oncogenes into oncogenes alone, e. g., Ras activation or Myc activation, is normally not sufficient to drive a cell into neoplasia. Excessive proliferative signals lead – via increased ARF expression, MDM2 inactivation by ARF and subsequent stabilization of p53 – to cell death, providing a safeguard against tumor development. Only when the pro-apoptotic functions of the network are inactivated at the same time will a cell be able to take advantage of its excessive proliferative capabilities, and the number of malignant cells will be ex-

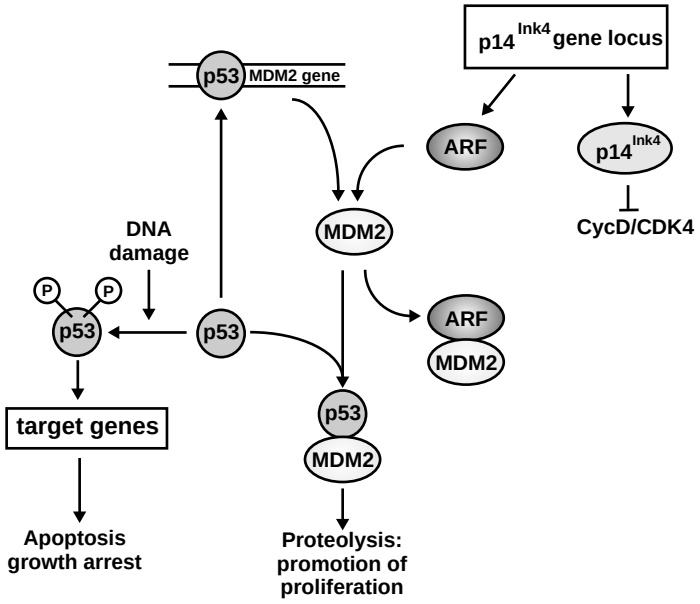


Fig. 14.15 The MDM2-p53 network. The level of p53 is strongly regulated by its interaction with MDM2. Binding of MDM2 to p53 targets p53 for proteolytic degradation thus keeping p53 concentration low. DNA damage induces phosphorylation of p53. The phosphorylated p53 is no longer bound by MDM2, proteolysis is decreased and the rising p53 levels initiate apoptosis and lead to growth arrest. A negative feedback loop exists between p53 and MDM2 since p53 controls the expression of MDM2 at the level of transcription. Another control is exerted by the tumor suppressor $p19^{ARF}$, which is coded for by the $p16^{Ink4}$ gene locus. $p19^{ARF}$ binds to MDM2 and sequesters it from the feedback loop thus helping to increase p53 concentration. The pro-apoptotic and anti-proliferative function of p53 may be impaired by mutation of p53, by overexpression of MDM2 or by a loss of $p19^{ARF}$ function. The $p16^{Ink4}$ gene locus is of specific importance in this network because it also codes for the inhibitor $p16^{Ink4}$ which inhibits the cyclinD/CDK4 complexes.

panded. Since p53 is also an essential component of DNA damage checkpoints, cells with a deficient p53-MDM2 network will be able to grow in the presence of DNA damage, allowing for increased genetic instability.

Inactivation of the p53 network can occur at many positions. Mutation of p53 itself is observed in the majority of tumors. These mutations are mostly nonsense mutations and affect predominantly the DNA-binding region of p53, impairing its transcriptional control function. Interestingly, mutated p53 proteins are often more stable than wild-type p53 and accumulate in cancer cells. The preferred mutation of the DNA-binding domain suggests that inactivation of the growth arrest and apoptotic functions of p53 is selected for during tumorigenesis

Loss of the p53 response can also occur via inactivation of other components of the MDM2-p53 network. MDM2 overexpression and amplification of the *mdm2* gene is observed in many tumors, leading to enhanced degradation of p53. Another way to

inactivate the p53-MDM2 module uses the ARF protein. A decreased level of ARF, which has the characteristics of a tumor suppressor, is frequently observed in tumor cells. This often occurs without mutation of the ARF gene itself. Rather, aberrant CpG-methylation of promotor sequences at the $p16^{INK4b}$ gene locus appears to be responsible in these cases for silencing the *ARF* gene.

Important features of cancer cells are enhanced genetic instability, dysregulated and autonomous cell proliferation, increased cell survival, and the ability to survive in a foreign cellular environment which exposes the cell to various stresses. Overall, these properties provide an advantage for the cancer cell.

The pathways that mediate genotoxic, non-genotoxic and oncogenic stress are the target of a multitude of selection pressures, and these pressures will converge on the p53-MDM2 module, acting either separately or in combination.

Inactivation of the p53 network enables the cell to continue in the cell cycle with damaged DNA, yet without DNA repair taking place. Furthermore, failure of the apoptotic control function permits the survival of cells with damaged DNA. Both effects lead to increased susceptibility of the genome to accumulation of further mutations. The cells can also divide under conditions in which serious changes of the genome are present, such as DNA amplification and chromosome rearrangement (see Kinzler and Vogelstein, 1996). Failure of the p53 function cancels a central control element that ensures the integrity of the genome. Therefore, p53 has been said to have the function of a "*guardian of the genome*".

14.9 The Tumor Suppressor APC and Wnt/β-Catenin Signaling

Inactivation of the gene for the tumor suppressor protein APC (adenomatous polyposis coli) has been found to be one of the early events during the development of most colorectal cancers. The *APC* gene encodes a multifunctional protein involved in central biological processes, including cell adhesion and migration, proliferation, apoptosis and differentiation (review: van Es et al., 2003). Most functions of the APC protein are linked to the Wnt/*β*-catenin signal transduction pathway (Fig. 14.16), which leads from the cell surface to the level of transcription. In this pathway, signals are transduced from the extracellular signaling protein Wnt to the transcriptional regulator *β*-catenin, which controls the expression of numerous target genes, including the gene for cyclin D1 and the transcription factor Myc. Transcriptional regulation is exerted in a heterodimeric complex composed of *β*-catenin and transcription factors of the TCF family (TCF, T-cell factor, review: Brantjes et al., 2002). The levels of *β*-catenin are regulated via phosphorylation by the Ser/Thr-specific protein kinase glycogen synthase kinase 3*β* (GSK3*β*), which is found in a complex with the APC protein, *β*-catenin, and the scaffolding protein Axin. In the absence of Wnt ligand, the protein kinase within this tetrameric complex is active, and *β*-catenin becomes phosphorylated and is thereby targeted for ubiquitin-dependent degradation.

Levels of *β*-catenin are high when the Wnt ligand is available for binding to its receptor, which belongs to the class of seven-helix transmembrane receptors. The

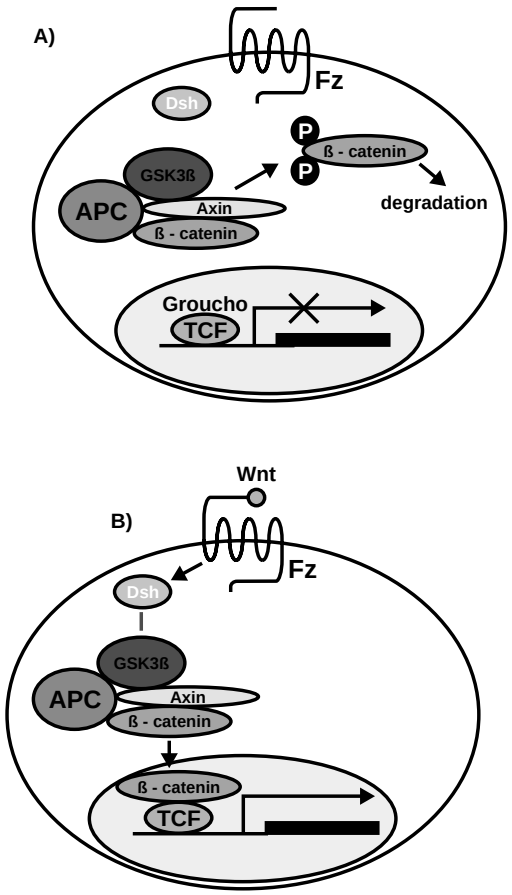


Fig. 14.16 Schematic overview of the Wnt signaling pathway. Central to wnt signaling is a protein complex composed of β-catenin, APC, glycogen synthase kinase 3 (GSK3) and axin.
A) In the absence of the extracellular signaling protein Wnt, β-catenin is phosphorylated by GSK3. Upon phosphorylation, β-catenin is primed for ubiquitinylation and subsequent degradation by the proteaosome. Reduced levels of β-catenin permit repression of Wnt target genes by association of transcriptional co-repressors like Groucho.
B) When a Wnt signal reaches a cell, Wnt associates with the transmembrane receptor Frizzled (FZ) leading to activation the cytoplasmic protein Dishevelled (Dsh) by an unknown mechanism. Dishevelled inhibits β-catenin phosphorylation and degradation and β-catenin now acts as transcriptional coactivator for the TCF family of transcription factors which activates target genes that regulate diverse cellular responses. Levels of β-catenin can be also regulated by p53 and by integrins.

Frizzled receptor is an example of receptors activated by the Wnt ligand. Upon activation of the Frizzled receptor, the protein kinase activity of GSK3β is inhibited, and β-catenin is stabilized and translocates to the nucleus, where it associates with the transcription factors of the TCF family. It is not known how the activated Frizzled receptor inhibits the protein kinase activity of GSK3β. The protein Disshelved is involved in this downregulation.

Loss of APC function leads to uncontrolled activation of the Wnt/β-catenin pathway and provides a proliferative advantage to the mutated cell (reviews:; Moon et al., 2002; Fodde, 2003). In addition, chromosomal instability is observed in cells with decreased APC function. The C-terminus of APC encompasses domains responsible for proper attachment of the mitotic spindle to the kinetochore, and it is assumed that loss of this function elicits chromosomal instability. It is remarkable that two central aspects of tumor formation, namely proliferative advantage and chromosomal instability, are observed upon deactivation of the Wnt/β-catenin pathway. Malfunction of other components of the Wnt/β-catenin pathway has also been observed in tumors. The scaffold protein axin has been

identified as a further tumor suppressor in this pathway, and β-catenin can be activated to an oncogene by mutation or aberrant expression. There is also a link to p53 function which downregulates β-catenin levels in response to DNA damage. Overall, aberrant function of the Wnt/β-catenin pathway is observed in nearly all colorectal cancers.

Reference

1 Alarcon-Vargas, D. and Ronai, Z. (2002) p53-Mdm2–the affair that never ends. *Carcinogenesis*, **23**, 541–547.
2 Blume-Jensen, P. and Hunter, T. (2001) Oncogenic kinase signalling. *Nature*, **411**, 355–365.
3 Brantjes, H., Barker, N., van Es, J. and Clevers, H. (2002) TCF: Lady Justice casting the final verdict on the outcome of Wnt signalling. *Biol.Chem.*, **383**, 255–261.
4 Cho, Y., Gorina, S., Jeffrey, P.D. and Pavletich, N.P. (1994) Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346–355.
5 Cohen, S.M. and Ellwein, L.B. (1990) Cell proliferation in carcinogenesis. *Science*, **249**, 1007–1011.
6 Fero, M.L., Randel, E., Gurley, K.E., Roberts, J.M. and Kemp, C.J. (1998) The murine gene p27Kip1 is haplo-insufficient for tumour suppression. *Nature*, **396**, 177–180.
7 Fodde, R. (2003) The multiple functions of tumour suppressors: it's all in APC. *Nat.Cell Biol.*, **5**, 190–192.
8 Giaccia, A.J. and Kastan, M.B. (1998) The complexity of p53 modulation: emerging patterns from divergent signals. *Genes Dev.*, **12**, 2973–2983.
9 Guimaraes, D.P. and Hainaut, P. (2002) TP53: a key gene in human cancer. *Biochimie*, **84**, 83–93.
10 Jackson, A.L. and Loeb, L.A. (1998) On the origin of multiple mutations in human cancers. *Semin.Cancer Biol.*, **8**, 421–429.
11 Jones, P.A. and Baylin, S.B. (2002) The fundamental role of epigenetic events in cancer. *Nat.Rev.Genet.*, **3**, 415–428.
12 Kinzler, K.W. and Vogelstein, B. (1997) Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature*, **386**, 761, 763.
13 Lengauer, C., Kinzler, K.W. and Vogelstein, B. (1998) Genetic instabilities in human cancers. *Nature*, **396**, 643–649.
14 Loeb, L.A. (1991) Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res.*, **51**, 3075–3079.
15 Loeb, L.A. (2001) A mutator phenotype in cancer. *Cancer Res.*, **61**, 3230–3239.
16 Look, A.T. (1997) Oncogenic transcription factors in the human acute leukemias. *Science*, **278**, 1059–1064.
17 Maru, Y. (2001) Molecular biology of chronic myeloid leukemia. *Int.J.Hematol.*, **73**, 308–322.
18 McCormick, F. (1999) Signalling networks that cause cancer. *Trends Cell Biol.*, **9**, M53-M56.
19 Menssen, A. and Hermeking, H. (2002) Characterization of the c-MYC-regulated transcriptome by SAGE: identification and analysis of c-MYC target genes. *Proc.Natl.Acad.Sci.U.S.A*, **99**, 6274–6279.
20 Michael, D. and Oren, M. (2002) The p53 and Mdm2 families in cancer. *Curr.Opin.Genet.Dev.*, **12**, 53–59.
21 Moon, R.T. and Miller, J.R. (1997) The APC tumor suppressor protein in development and cancer. *Trends Genet.*, **13**, 256–258.
22 Ru, H.Y., Chen, R.L., Lu, W.C. and Chen, J.H. (2002) hBUB1 defects in leukemia and lymphoma cells. *Oncogene*, **21**, 4673–4679.
23 van Es, J.H., Barker, N. and Clevers, H. (2003) You Wnt some, you lose some: oncogenes in the Wnt signaling pathway. *Curr.Opin.Genet.Dev.*, **13**, 28–33.
24 Yarden, R.I., Pardo-Reoyo, S., Sgagias, M., Cowan, K.H. and Brody, L.C. (2002) BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage. *Nat.Genet.*, **30**, 285–289.
25 Zheng, L., Annab, L.A., Afshari, C.A., Lee, W.H. and Boyer, T.G. (2001) BRCA1 mediates ligand-independent transcriptional repression of the estrogen receptor. *Proc.Natl.Acad.Sci.U.S.A*, **98**, 9587–9592.

15
Apoptosis

Apoptosis is a naturally occurring process by which a cell is directed to programmed death. The name apoptosis was coined following investigations of the nematode *Caenorhabditis elegans* and is of Greek origin, describing the falling of leaves. The course of apoptosis is accompanied by characteristic changes in cell morphology, shown schematically in Fig. 15.1. Condensation of the chromatin, degradation of DNA, cell shrinkage, fragmentation of the cell nucleus, and disassembly into membrane-enclosed apoptotic vesicles are characteristics that clearly distinguish apoptosis from another form of cell death known as necrotic cell death.

15.1
Basic Functions of Apoptosis

Apoptosis is based on a genetic program that is an indispensable part of the development and function of an organism. It serves to eliminate undesired or superfluous cells in a targeted manner. The conditions under which the apoptotic program is activated are very diverse:

- **Tissue homeostasis**
  Apoptosis is considered to be of central importance in homeostasis of tissues: in an organ or a tissue, the cell number must be kept constant within narrow limits. An increase in cells due to cell division is compensated for by processes to eliminate cells that are no longer functional or are old. Apoptosis is a process that helps to keep the cell number in a tissue within limits that are suitable for the development and function of the organism. In this process, it is an important function of programmed cell death to eliminate cells that have a defect in signaling pathways regulating cell proliferation. Cells that have acquired an oncogenic potential due to the aberrant activation of proto-oncogenes can be eliminated by apoptosis. If defects occur in the apoptotic program, the consequence may be a pathological increase or decrease in the number of cells (Fig. 15.2). Examples of diseases associated with an increased rate of cell survival are cancer and autoimmune diseases. Diseases associated with increased apoptosis include AIDS and neurodegenerative diseases (Thompson, 1995).

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

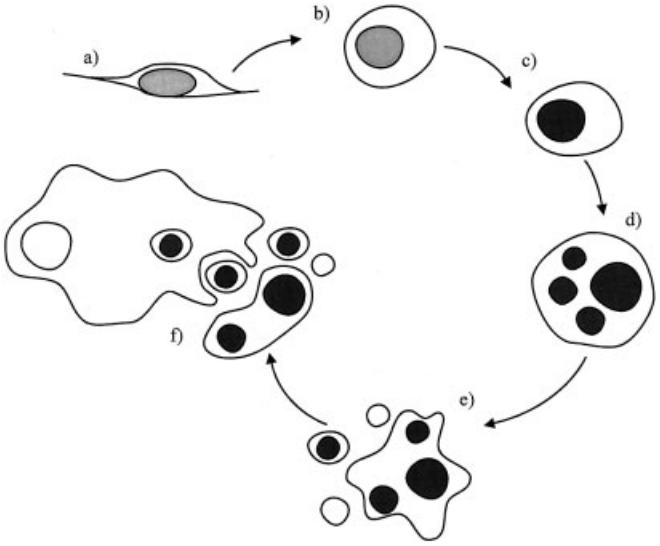


Fig. 15.1 Schematic representation of morphologic changes in a cell during apoptosis. On reception of an apoptotic signal, an adhesive cell (a) begins to become rounded (b) and the nuclear DNA condenses (c). The DNA is fragmented and the nucleus begins to break down into discrete chromatin bodies (d). Finally, the cell disintegrates into several vesicles (apoptotic bodies) (e), which are phagocytozed by neighboring cells (f).

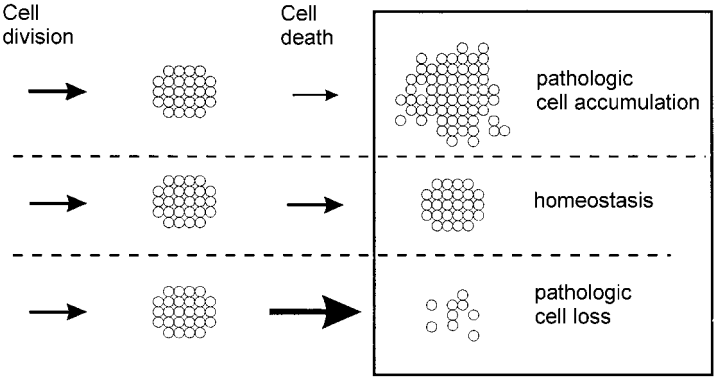


Fig. 15.2 Influence of apoptosis on homeostasis of a cell grouping. In a grown organism, the cell number in a tissue is determined by the relation between the rate of cell division and cell death. The rates of both processes are represented in the figure by the size of the arrow. In a normal tissue, the cell number remains constant (homeostasis) since both processes occur at the same rate. If the rate of cell proliferation predominates, diseases occur characterized by increased cell number (e. g., in tumors). In the reverse case, when the rate of cell death predominates, the cell number is reduced in a pathologic fashion. In the absence of compensatory changes in the cell division rate, changes in the extent of apoptosis can lead to either accumulation of cells or loss of cells. According to Thompson, (1995).

- **Development and differentiation**
  Apotosis has an indispensable role in development and differentiation processes (review: Vaux and Korsmeyer, 1999), especially in the embryo. Here, it provides a means to switch off cells no longer needed during embryonal morphogenesis and synaptogenesis. The genetic description of programmed cell death originates from observations on the development of the nematode *Caenorhabditis elegans*, where 1090 cells are formed of which 131 are eliminated in a targeted manner by an intrinsic apoptotic program.
- **Immune system**
  In the immune system, T- and B-cells undergo apoptosis in many instances of their development, homeostasis, and activation (review: Krammer, 2000; Rathmell and Thompson, 2002). Examples include
  – elimination of target cells (e. g., virus-infected cells) by cytotoxic T-lymphocytes
  – elimination of autoreactive B- or T-lymphocytes, natural selection and elimination of cells in the thymus and bone marrow: the vast majority of T-cells that migrate to the thymus are eliminated by apoptosis.
- **Cell damage**
  Another function of apoptosis is the destruction of damaged cells. The apoptotic program may be activated in the presence of cell damage or during stress (review: Evan and Littlewood, 1998). Cells with damaged DNA can be eliminated with the help of apoptotic programs before they have the chance to accumulate mutations and possibly degenerate into a tumor cell.

15.2 Overview of Apoptosis

Starting from genetic studies on apoptosis in *C. elegans*, the pathways and components of apoptosis in mammalian cells have now been largely identified. At the center of the apoptotic program is a family of proteases named caspases. The caspases are involved in the initiation and execution of the program and can be activated by a large number of stimuli via two central pathways, one involving mitochondria, the other using transmembrane receptors of the tumor necrosis factor α (TNFα) class. Activation of apoptosis via mitochondria is an intrinsic pathway where stress signals, DNA damage signals, and defects in signaling pathways are processed. The TNFα (or death receptor) pathway is an extrinsic pathway that uses external signaling proteins for activation of the apoptotic program, and this pathway is mainly used in developmental processes and in the immune system.

As a consequence of caspase activation, a number of key enzymes and structural proteins of the cell are degraded, leading to cell death. The stimuli that induce apoptosis are very diverse and include, e. g., DNA damage, stress conditions, and malfunction of pathways regulating cell proliferation. In the normal situation of a tissue, a finely tuned balance exists between pro-apoptotic signals that activate the apoptotic program and anti-apoptotic signals that suppress apoptosis and promote cell survival. The homeostasis achieved by this balance can be disturbed in favor of apoptosis

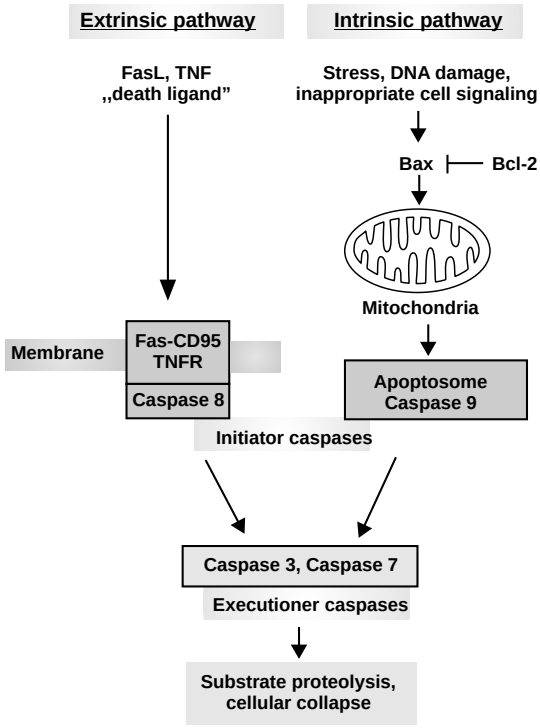


Fig. 15.3 The major pathways of apoptosis. The extrinsic pathway uses extracellular death ligands (Fas ligand, tumor necrosis factor (TNF)) to activate 'death receptors' which pass the apoptotic signal to initiator caspases (e. g. capsase 8) and to the executioner caspases (e. g. caspase 3; caspase 7). In the execution phase of apoptosis, various cellular substrates are degraded leading to cellular collapse. The intrinsic pathway uses the mitochondria as a central component for activation of apoptosis. In this pathway, a multitude of intracellular signals including various stresses, DNA damage and inappropriate cell signaling lead to activation of the pro-apoptotic protein Bax which induces release of cytochrome c from mitochindria, formation of the apoptosome and activation of the initiator caspase 9. Finally, the executioner caspases are activated and cells are destructed by proteolysis. Apoptosis via this pathway can be controlled by various antiapoptotic proteins including the Bcl-2 protein and inhibitors of apoptosis.

by a lack of survival signals (e. g. growth factors) or by a surplus of pro-apoptotic signals. Furthermore, a defect in the apoptotic program or excessive proliferation signals will enhance cell survival.

The major part of the apoptotic program exists in the cell in a latent, inactive form, and it only requires an apoptotic stimulus to activate the program and to initiate apoptosis. Thus, apoptotic processes can also take place without activation of transcription. There are also forms of apoptosis that are dependent on transcription.

An overview of apoptosis is shown in Fig. 15.3. The function and regulation of the components of apoptosis are discussed in more detail below.

15.3 Caspases: Death by Proteolysis

A family of specialized proteases, named caspases, is central to the apoptotic program (review: Grütter, 2000). The name caspase is a contraction of cysteine-dependent, aspartate-specific protease. These proteases use a Cys residue as a nucleophile and cleave the substrate after an Asp residue. The only known eukaryotic proteases with this specificity are the caspases themselves and the cytotoxic serine protease granzyme B from T-lymphocytes.

To date, 14 mammalian caspase sequences (named caspases 1 – 14) have been reported, of which eleven are of human origin. With respect to function, caspases are grouped into two biologically distinct subfamilies. One subfamily mediates initiation (initiator caspases, caspases-8, -9) or execution (executioner or effector caspases, caspases-3, -6 and -7) of the apoptotic program. Members of the other subfamily (caspases-1, -4, -5, -11, -12, and -14) are involved in inflammatory processes by processing pro-inflammatory cytokines.

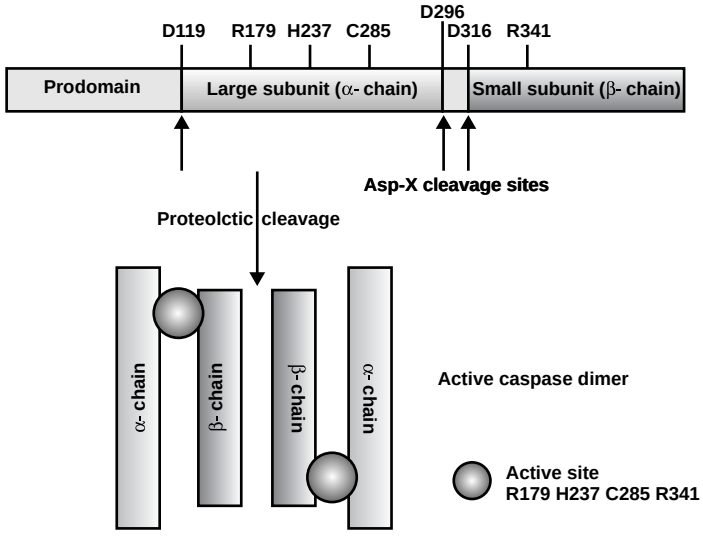
Structure and Mechanism

Like many other proteases, the caspases are formed as inactive proenzymes of 30 – 50 kDa and are activated by proteolytic processing. The proenzymes have an N-terminal prodomain followed by sequences encoding first a large (α-chain) and then a small (β-chain) subunit. The large and small subunits of the caspase are formed from the proenzyme by proteolysis at Asp-X sites. These internal cleavage sites are consistent with the ability to autoactivate or to activate other caspases as part of an amplification cascade.

Structural studies have shown that caspases form tetramers with two active sites. The catalytically active subunit of a caspase is made up of one α-subunit (17-12 kDa) and one β-subunit (10-13 kDa), which form a heterodimer with an active site composed of residues from both subunits. Two heterodimers then align to form a tetramer with two catalytic centers (Fig. 15.4). Essential residues of the active site comprise Arg179 and Arg341 (numbering for caspase-3) for binding the aspartic residue proximal to the cleavage site, His237, as part of a protease charge relay, and Cys285 as a nucleophile. Hydrogen bond interactions between active site residues of caspase 8 and a peptide inhibitor Z-DEVD-CHO are shown in Fig. 15.5.

The cleavage mechanism of the caspases is depicted schematically in Fig. 15.6. A typical protease mechanism is used, with a catalytic dyad for cleavage of the peptide bond. The nucleophilic thiol of Cys285 forms a covalent thioacyl bond to the substrate during the catalysis. The imidazole ring of His237 is also involved in the catalysis, facilitating hydrolysis of the amide bond by acid/base catalysis.

The special feature of the caspases is their high cleavage specificity. Recognition of the substrate occurs predominantly in a cleft formed by loop regions of the α- and β-



Caspase families: cleavage specificity

Group I	Group II	Group III
Caspase- 1,- 4,- 5,- 13	Caspase- 2,- 3,- 7	Caspase- 6,- 8,- 9,-10
WEHD↓	DEXD↓	(I/V/L)EXD↓

Fig. 15.4 Processing and subunit structure of caspases. Schematic representation of the proteolytic activation of caspases. Caspases are synthesized as single-chain precursors. Activation proceeds by cleavage of the N-terminal peptide at Asp119 and at the conserved sites Asp296 and Asp316 (all caspase-1 numbering convention), leading to the formation of a large α-subunit and a small β-subunit. The activity- and specificity-determining residues R179, H237, C285 and R341 are brought into the characteristic structural arrangement for catalysis. C285 is the catalytic nucleophile and H237 represents the general base. The crystal structures reveal that the active enzyme is a dimer in which one ($\alpha\beta$) unit that harbors the active site is related by a twofold axis to a second unit to form the active $(\alpha\beta)_2$ dimer. By their celavage specificity, caspases are grouped in three families. Caspases recognize a sequence of four amino acids and cleave the substrate C-terminal to the aspartate residue as indicated by the arrow. After Gruetter (2000).

subunits. The cleft recognizes a tetrapeptide located N-terminally to the canonical cleavage site Asp-X. The various caspases cleave different substrates. By virtue of cleavage specificity, the caspases can be grouped into three families, which differ mainly in position P4 of the tetrapeptide recognized on the substrate protein (see Fig. 15.4B).

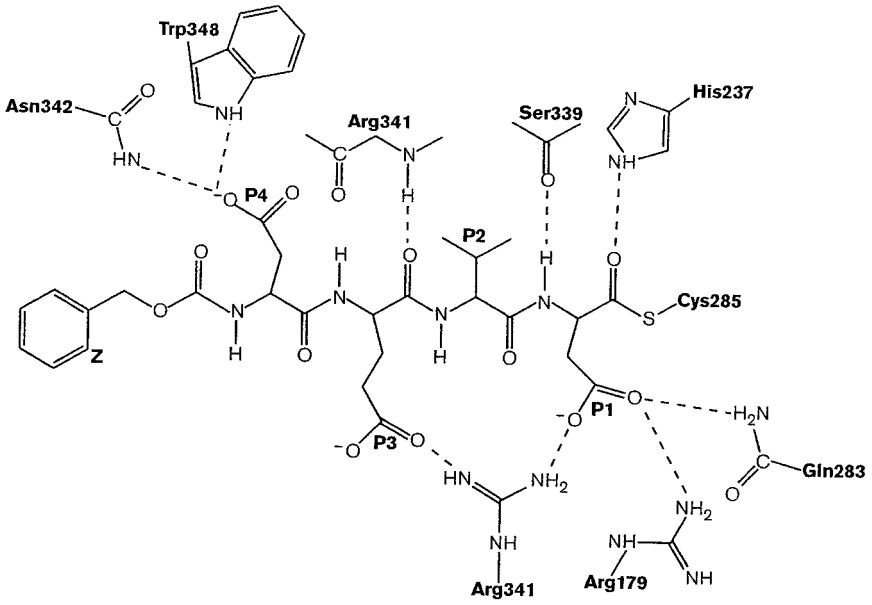


Fig. 15.5 Caspase-inhibitor interactions. Schematic drawing of the hydrogen-bond interactions in caspase–inhibitor complexes, exemplified by the caspase-8–benzyl(Z)-DEVD-CHO structure. Hydrogen-bond interactions are depicted as dashed lines. Caspase-1 numbering convention. After Gruetter (2000).

Substrates

A large number of caspase substrates have been identified, some of which have a direct relationship to the survival of the cell. The caspase substrates can be grouped into different classes according to their function:

1. **Procaspases**
   Triggering of caspase cascades involves the transactivation of a procaspase by already activated caspases. Thereby, sufficient proteolytic activity is generated to overwhelm endogenous caspase inhibitors, e. g., an inhibitor of apoptosis proteins (IAPs).
2. **Pro- and anti-apoptotic proteins**
   Examples of anti-apoptotic proteins degraded by caspases are the Bcl-2 and Bcl-x_L proteins, which are cleaved by caspase-3 to generate C-terminal fragments that are pro-apoptotic. Caspase-8 cleaves the pro-apoptotic protein Bid (see Section 15.4), generating a C-terminal fragment that induces release of cytochrome c from mitochondria.
3. **DNAase inhibitor ICAD (inhibitor of caspase-activated DNAase)**
   Caspase-mediated degradation of ICAD relieves inhibition of a DNAase responsible for DNA fragmentation (Enari et al, 1998).

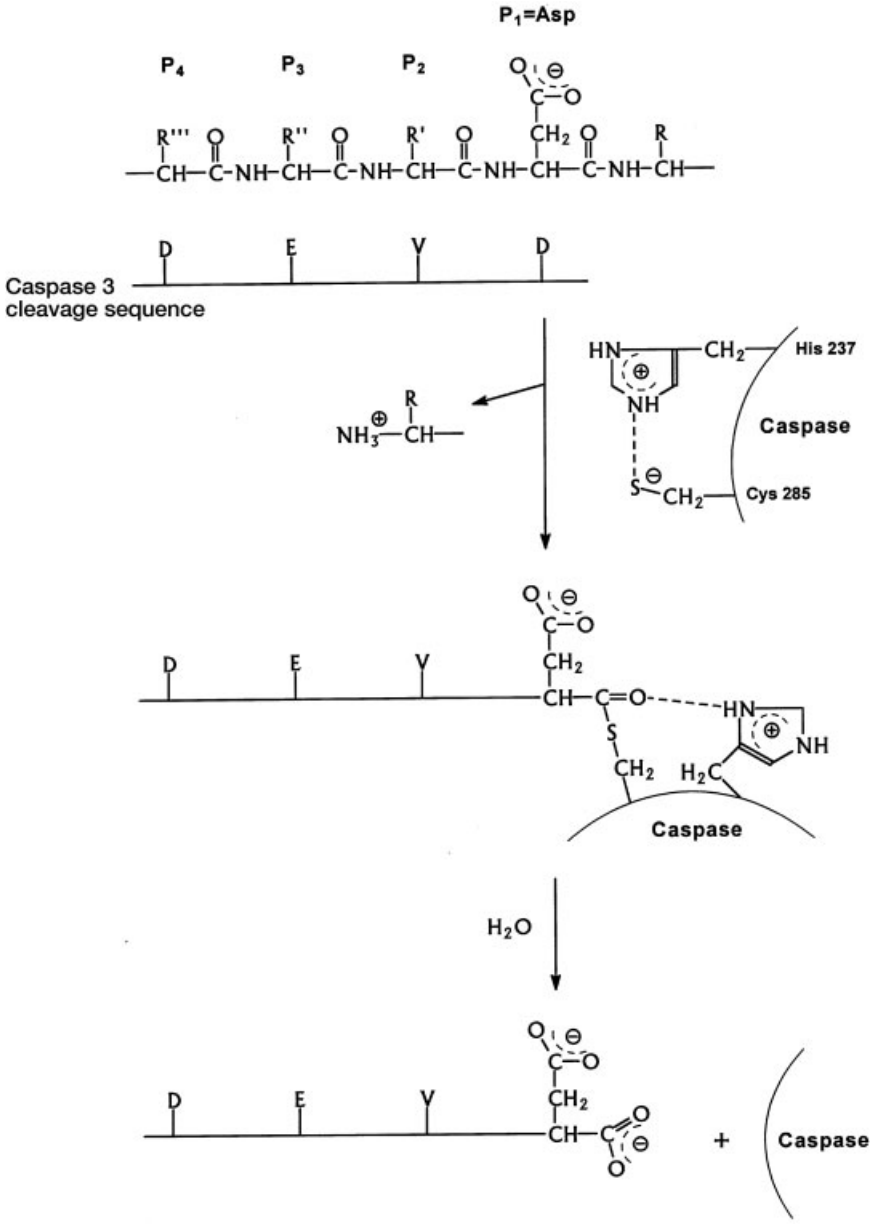


Fig. 15.6 Substrate recognition and postulated cleavage mechanism of caspases. For caspase 3, four specific residues N-terminal to the cleavage site are required for cleavage in addition to the essential Asp residue. In the first step of the reaction, a covalent thioacyl-intermediate is formed between the N-terminal part of the substrate and the caspase, this is hydrolytically cleaved in the second step.

4. **Structural proteins**
   Gelsolin and lamin are substrates whose degradation is responsible for part of the subcellular structural changes observed during apoptosis.
5. **Proteins important for cellular signaling, DNA repair and macromolecular synthesis**
   Examples include the focal adhesion kinase FAK (see Section 11.3), β-catenin (see Section 14.9), p21-activated kinase (PAK), replication factor C, and poly-ADP-ribose polymerase. The latter enzyme participates in the repair of DNA double-strand breaks.

Activation and Regulation

Unprogrammed activation of the caspases has serious consequences for the cell. Therefore, activation of caspases is strictly controlled. In the normal state of the cell, the caspases are maintained in an inactive state but can be rapidly and extensively activated by a small inducing signal. Control of caspase activity occurs at two levels. The first level of caspase regulation involves the conversion of the caspase precursors, the zymogens, to the active forms in response to inflammatory or apoptotic stimuli. The second level of caspase control involves the specific inhibition by binding of natural inhibitors.

Control by Proteolytic Activation

Proteolytic activation of procaspases uses two mechanisms, depending on whether the caspase functions as an effector caspase (caspases-3, -6 and -7) or as an initiator caspase (caspases-8 and -9). Initiator caspases are the first to be activated in response to a proapoptotic stimulus and are responsible for activating the effector caspases by limited proteolysis. The effector caspases are thought to be responsible for most of the substrate proteolysis observed during apoptosis. By digesting central proteins, the effector caspases direct the cell to death.

The initiator caspases receive proapoptotic signals and initiate the activation of a caspase cascade. They are activated by assembly into a multiprotein complex, and they contain large prodomains responsible for this interaction.

The two main pathways of apoptosis, the death receptor pathway and the mitochondrial pathway, differ in the mechanism of activation of the initiator caspase but use the same effector caspase at least partially.

The cell uses two mechanisms for proteolytic activation of caspases:

- **Proteolytic cleavage in a caspase cascade**
  A procaspase is proteolytically activated by a caspase already activated upstream in the apoptotic signaling pathway. This mechanism is used particularly in the effector part of the apoptotic pathway.
- **Aggregation-induced activation of an initiator procaspase**
  In this pathway, an apoptotic stimulus induces assembly of the procaspase into a multiprotein complex. Activation is assumed to occur by autoproteolyis involving mutual cleavage in trans within this complex. Specific cofactors like FADD (see Section 15.6) and Apaf1/Cytc (see Section 15.5) are involved in the formation of

the multiprotein complex. The mechanism underlying activation in the multiprotein complexes is not well understood. The model of "proximity-induced activation" states that an initiator procaspase of low activity is activated by stimulus-induced aggregation (review: Hengartner, 1998). The procaspases have low protease activity corresponding to only 1–2% of that of the mature caspase. If a procaspase is recruited into a larger aggregate, this activity is thought to be sufficient for mutual processing of the procaspase molecules to mature caspases.

Interaction Modules in Caspase Multiprotein Complexes

Recruitment of the initiator procaspases into a multiprotein complex results from a regulated series of protein-protein interactions mediated by "interaction modules". Four types of "interaction modules" are involved in the activation of initiator caspases and thus play important roles in the initiation of apoptosis (review: Weber and Vincenz, 2001). These domains have been named the death domain (DD), , the death effector domain (DED), the caspase activation and recruitment domain (CARD), and the less characterized pyrin domain. The domains are found on several components of the apoptotic signaling pathways and mediate homotypic protein-protein interactions, i. e., a given module will interact only with a member of the same family and not with members of the other families. Since members of the same module are found on different proteins, these modules mediate the assembly of hetero-oligomeric protein complexes. As examples, DDs are found on death receptors and their cofactors, DEDs on cofactors and the initiator caspase-8, and CARDs on cofactors, caspase-2, and caspase-9.

Control by Inhibitor Proteins

Caspases can be directly inhibited by binding of inhibitory proteins. We know of three families of proteins capable of ablating caspase activity *in vivo* and *in vitro* (review: Stennicke et al., 2002). One of these, the inhibitor of apoptosis (IAP) family, is conserved from flies to man and regulates cellular apoptosis by direct caspase inhibition. The two other types of inhibitors are of viral origin. Structural analysis of inhibitor-caspase complexes has shown that the IAPs inhibit caspases by obstructing the binding sites responsible for the recognition of the P4-P1 substrate residues. Overall, the mechanism of inhibition by IAPs is different from that of other protease inhibitors. The IAPs themselves are subject to regulatory influences and are part of signaling pathways that link apoptotic stimuli to caspase activity.

15.4 The Family of Bcl-2 Proteins: Gatekeepers of Apoptosis

Bcl-2 family members are the key regulators that control mitochondria-mediated apoptosis (review: Parone and Martinou, 2002). The founding member of this protein family, the Bcl-2 protein, was first identified as an oncoprotein coded by a gene affected by translocations of chromosomes 14 and 18 in B cell lymphomas. It was soon shown,

however, that the Bcl-2 protein is not involved in the regulation of the cell cycle, in contrast to many other oncoproteins, and thus does not fit into the classical oncogene picture. Furthermore, homology was established with the Ced9 protein of *C. elegans*, which has an antiapoptotic function in this organism.

At present, >20 members are known of the *Bcl-2 family*, which can have a negative or a positive effect on the initiation of the apoptotic program. All Bcl-2 family members have at least one copy of a so-called BH motif (BH, Bcl-2 homolog), of which there are four types (BH1 – BH4). On the basis of structural and functional criteria, the Bcl-2 family has been divided into three groups (review: Cory and Adams, 2002; Fig 15.7):

- **Group I** contains proteins that harbor BH domains 1 – 4 and a hydrophobic C-terminal tail with which they span the cytosolic surface of various intracellular membranes, such as the outer mitochondrial membrane. All members of this group have anti-apoptotic functions. It is assumed that the members of this family prevent cell death by binding and sequestering the pro-apoptotic Bcl-2 family members of groups II and III. Overexpression of group I members can prevent initiation of the apoptotic program in various cell types. The oncogenic function of Bcl-2 protein, observed in association with its overexpression, can be explained by its antiapoptotic effect: the high level of Bcl-2 protein suppresses initiation of the apoptotic program and an important requirement for further tumor progression is fulfilled. In this situation, damaged cells, which would have been eliminated by apoptosis in the normal situation, can survive.
- **Group II** includes the pro-apoptotic Bax and Bak proteins. These are similar in structure to the group I members but lack the N-terminal BH4 domain. Their activity is necessary for the induction of mitochondria-mediated apoptosis.
- **Group III** consists of large proteins that contain a single BH3 domain. They act by binding to group I and/or group II family members via their BH3 domain. These

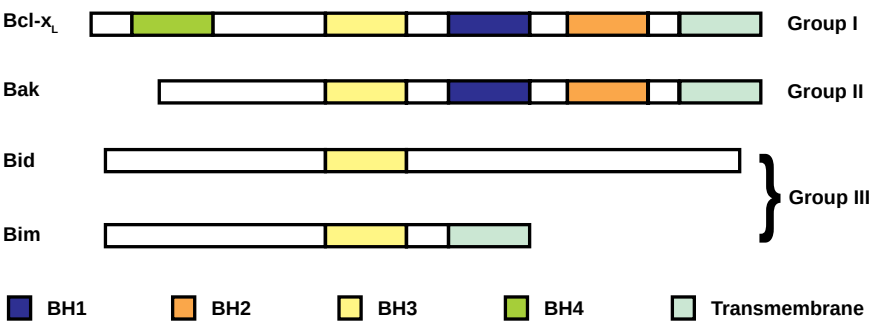


Fig. 15.7 Domain structure of the Bcl-2 family. On the basis of functional and structural criteria, the Bcl-2 family has been divided into three groups. Group I comprises anti-apoptotic proteins characterized by four short , conserved Bcl-2 homology (BH) domains, known as BH1-BH4. Group II includes the pro-apoptotic proteins Bax and Bak which are similar in structure to the Group I proteins but lack the N-terminal BH domain. Group III consists of pro-apoptotic polypeptides including Bid, Bad, Nora etc. that contain a single BH3 domain.

pro-apoptotic polypeptides include the Bid, Bad, Bim, Noxa and Puma proteins, which are thought to be sensors of diverse pro- and anti-apoptotic stimuli, integrating these into a life-or-death decision.

15.5 The Mitochondrial Pathway of Apoptosis

Apoptosis is initiated by two principal pathways. The mitochondrial pathway is intrinsic and is activated from within the cell. The other pathway (see Section 15.6) is of an extrinsic nature and is initiated by external ligands that bind to and activate transmembrane receptors called death receptors.

The mitochondria have emerged as a central component of the intrinsic apoptotic signaling pathways and are now known to control apoptosis via the release of apoptogenic proteins (Fig.15.8). The apoptotic signals that are channeled through the mitochondrial pathway of apoptosis include various stresses like DNA damage, oxidative stress, UV radiation, protein kinase inhibition, and growth factor deprivation.

Cytochrome c Release

Apoptotic stimuli are relayed to mitochondria via the pro-apoptotic proteins of the Bax/Bak group, which translocate, upon receipt of the appropriate stimuli, from the cytosol to the outer membrane of the mitochondria and thereby allow the release of different mitochondrial proteins that are normally present in the intermembrane space of these organelles. Of these proteins, cytochrome c is most important for triggering further downstream apoptotic events. In the unstimulated situation, cytochrome c is trapped at the inner mitochondrial membrane by interaction with the lipid cardiolipin. Furthermore, the pro-apoptotic Bax/Bak proteins are kept in check by heterodimerization with the anti-apoptotic Bcl-2/Bcl-XL proteins. The mechanisms which initiate the release of cytochrome c upon receipt of an apoptotic stimulus remain elusive. The driving force for transport of the Bax/Bak proteins to the outer mitochondrial membrane is unclear, and it is not well understood how cytochrome c is released. According to one model of the release of cytochrome c, oligomerization of Bax/Bak proteins leads to the formation of a pore in the outer mitochondrial membrane through which cytochrome c can translocate into the cytosol and subsequently become integrated into the apoptosome.

In addition to cytochrome c, other soluble proteins contained in the intermembrane space of mitochondria are released through the outer membrane and participate in the organized destruction of the cell. Among these proteins are the apoptosis-inducing factor (AIF), endonuclease G, and the Smac/Diablo proteins. The AIF protein has been shown to be responsible for triggering a pathway of programmed cell death that is distinct from the caspase pathways. Endonuclease G translocates during apoptosis from the mitochondrion to the nucleus, where it cleaves chromatin DNA into nucleosomal fragments independently of caspases. The Smac/Diablo proteins are known to counteract the function of the caspase inhibitors, the IAPs (see above).

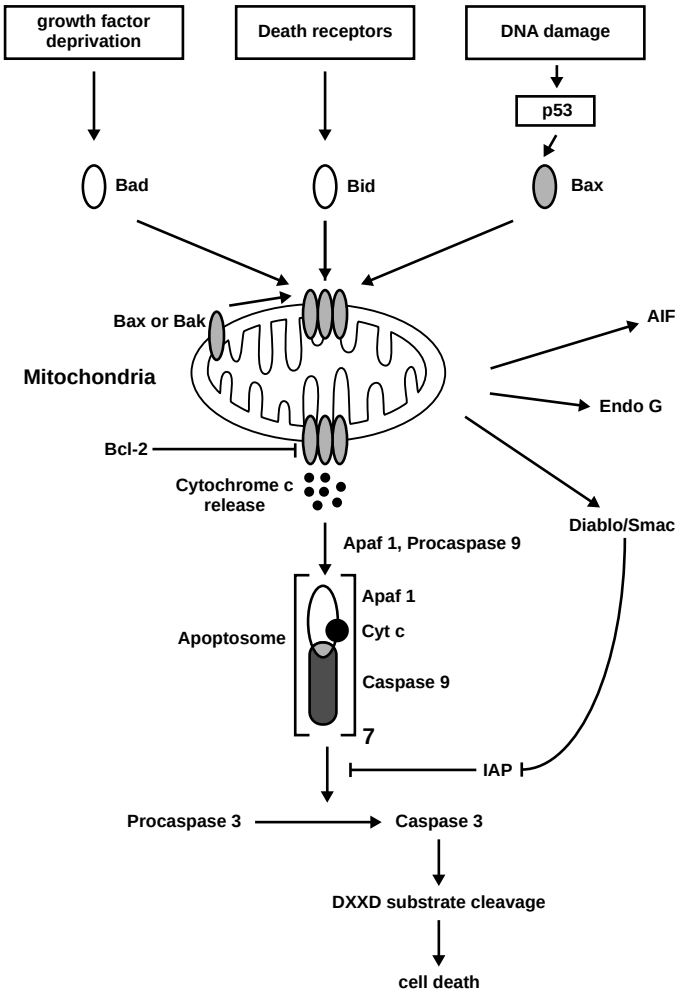


Fig. 15.8 The mitochondrial pathway of apoptosis. Cellular stress (e. g. growth factor deprivation, activation of death receptors, DNA damage) promotes the release of cytochrome c from mitochondria in a process involving death-promoting members of the Bcl-2 family (e. g. Bid, Bax, Bad, Bak). These proteins are assumed to translocate to the mitochondria or to undergo conformational changes within the outer mitochondrial membrane forming a pore-like structure which facilitates escape of cytochrome c from the mitochondrium. Cytochrome c assembles with Apaf1 and procaspase 9 to form the apoptosome which is composed of seven procaspase 9/Apaf1/cyt c trimers. The initiator procaspase 9 is activated in this complex and triggers the execution phase of apoptosis leading finally to cell death. Negative regulation of the mitochondrial pathway occurrs at the level of cytochrome c release and caspase activity. Cytochrome c release can be blocked by anti-apoptotic proteins like Bcl-2. Mature caspases are subject to inhibition by the conserved Inhibitors of apoptosis (IAP) family of proteins. Other proteins released by mitochondria comprise the apoptosis-inducing factor (AIF), endonuclease G (Endo G) and the Smac/Diablo proteins. The latter family of proteins interferes with the IAP function.

Formation of the Apoptosome and Triggering of a Caspase Cascade

In a further step of apoptosis, the cytochrome c released from the mitochondria promotes the assembly of a multiprotein complex, termed apoptosome, which contains cytochrome c, the adaptor protein Apaf1, and procaspase-9. The apoptosome requires ATP for its formation and is able to cleave and activate procaspase-3, an effector caspase. The adaptor protein Apaf1 appears to play a major structural role in this assembly. Apaf1 contains WD motifs for interaction with cytochrome c and a CARD motif, which directs binding to the CARD motifs of procaspase-9 and procaspase-3. Structural studies on the apoptosome by electron microscopy have revealed a wheel shaped heptameric complex, with the CARD domains of Apaf1 located at the central hub and the WD40 repeats at the extended spokes (Acehan et al., 2002). The location of procaspase 9 in this complex is still open as is the mechanism of caspase 9 activation.

The activation of caspase-9 by apoptosome formation sets in motion a cascade of caspase activation events. At the top of this cascade is caspase 3, which cleaves other downstream effector procaspases (caspases-2, -6, -8 and -10) or apoptotic substrates containing the recognition motif DXXD. Activation of this hierarchically structured cascade leads to proteolyis of multiple substrates, and the cell is committed to death. The cellular infrastructure is destroyed, and changes at the plasma membrane are triggered that promote engulfment by phagocytes.

15.6 Death Receptor-triggered Apoptosis

One major pathway of apoptosis is activated by external ligands that bind to and activate receptor systems known as death receptors (Fig. 15.9). The death receptors are transmembrane receptors that belong to the superfamily of tumor necrosis factor (TNF) receptors. Members of the death receptor family are characterized by a Cys-rich extracellular domain and a homologous intracellular domain known as the "death domain", DD. The death receptor family includes a receptor termed Fas (also known as CD95), tumor necrosis factor α (TNF-α) receptor 1 (TNF-R1), and two other receptors, DR4 and DR 5. In the absence of ligands, these receptors exist as inactive trimeric complexes. The ligands of these receptors are trimeric extracellular proteins that bind to the extracellular portion of the trimeric receptor and thereby activate the receptor for transmission of a signal into the interior of the cell. Depending on the cellular context, the death receptors can transmit pro-apoptotic, anti-apoptotic, anti-inflammatary, or pro-inflammatory signals. In the following, the main features of the pro-apoptotic signaling pathways mediated by Fas/CD95 and TNF-R1 will be presented. For details of these and related pathways, the reader is referred to Baud and Karin (2001), Rathmell and Thompson, (2002), Chen and Goeddel (2002).

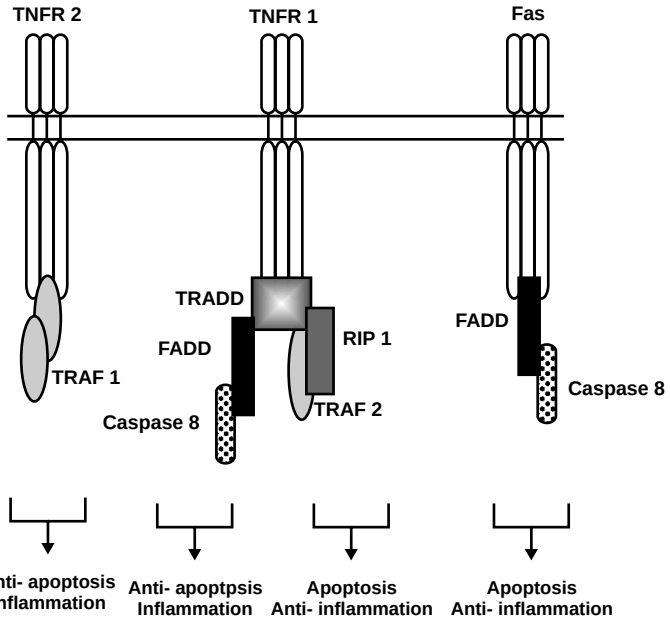


Fig. 15.9 Proximal components of the tumor necrosis factor (TNF) receptor type 1 (TNFR1), TNF receptor type 2 (TNFR2) and Fas signal-transduction pathways and their relationships to the activation and inhibition of programmed cell death and inflammation. Both TNFR1 and Fas transduce apoptotic and anti-inflammatory signals through the recruitment of Fas-associated death domain protein (FADD) and subsequent recruitment and activation of caspase 8. TNFR1 also mediates anti-apoptotic and inflammatory responses through the recruitment of TNF-receptor-associated factor 2 (TRAF2) and receptor-interacting protein 1 (RIP1). Other members of the TNF receptor family, such as TNFR2, recruit TRAF2 and TRAF1 to transmit their anti-apoptotic and inflammatory signals. TRADD: TNFR1-associated death domain protein (after Baud and Karin, 2001).

15.6.1
The Fas/CD95 Signaling Pathway

Fas/CD95 has a central role in the physiological regulation of programmed cell death in the immune system, where it is mainly used to instruct lymphocytes to die during immune responses. A deficieny in Fas/CD95 can result in abnormal lymphoid development and autoimmune diseases (review: Krammer, 2000).

The ligand for the Fas/CD95 receptor (Fas ligand/CD95 ligand) is a homotrimeric protein that binds to Fas/CD95, causes clustering and activation of inactive Fas/CD95 complexes, and allows the formation of a death-inducing signaling complex (DISC). The Fas-DISC (Fig. 15.10) contains the adaptor protein Fas-associated death domain protein (FADD) and caspases-8 or -10, which can initiate the process of apoptosis. Clustering of the components of Fas-DISC is mediated by homotypic interactions between death domains found on Fas/CD95 and on FADD, and between death effec-

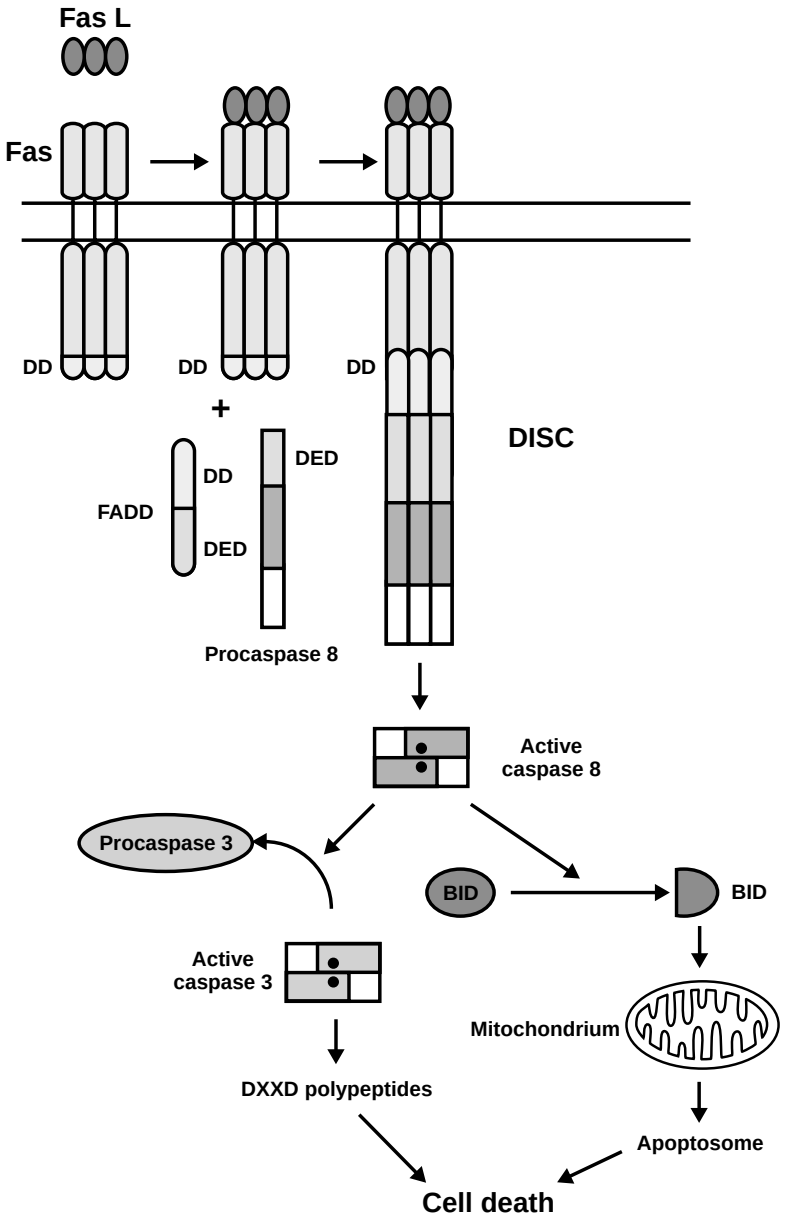


Fig. 15.10 The Fas signaling pathway. Binding of Fas ligand to Fas/CD95 triggers formation of the death-inducing signaling complex (DISC) composed of Fas ligand, Fas, FADD and procaspase 8. The latter is activated in the DISC to form the mature caspase 8 which can transduce the apoptotic signal by two ways. In one way, caspase 8 produces mature effector caspase 3 from its precursor leading to proteolysis of subtrates containing the DXXD motif. In another reaction, caspase 8 cleaves the Bcl-2 family protein Bid whose truncated form initiates the mitochondrial pathway of apoptosis by triggering cytochrome c release and apoptosome formation. FADD: Fas-associated death domain protein; DD: death domain; DED: death effector domain.

tor domains found on FADD and procaspase-8. As a result of Fas ligand-induced clustering of Fas/CD95, FADD, and caspase-8 or caspase-10, these initiator caspases are processed in an autoproteolytic way by induced proximity. The processed, active caspases-8 or -10 are then released from the DISC and activate downstream apoptotic proteins.

Depending on the cell type, two different downstream pathways are triggered. In type I cells, processed caspase-8 produced in large amounts directly activates a caspase cascade. Among the caspases activated are caspase-3, which cleaves other caspases or vital substrates of the cell and thus paves the way for the execution phase of apoptosis. In type II cells, proper activation of effector caspases requires amplification via the mitochondrial pathway of apoptosis. Here, smaller amounts of active caspase-3 are produced which cleave the pro-apoptotic Bcl-2 family member Bid. The truncated form of Bid "activates" mitochondria by an unknown mechanism, which now release pro-apoptotic proteins like cytochrome c and Smac/Diablo (see Section 15.5). Cytochrome c release triggers the formation of the apoptosome, resulting in the activation of caspase-9 and subsequently caspase-3, which in turn can activate caspase-8 outside the Fas-DISC.

Procaspase-8 functions as an initiator caspase in this system, since its activation is the signal for activation of the downstream caspase cascade. The DED motif of caspase-8 is localized on its large prodomain. Similar motifs are found in other caspases with large prodomains (caspases-2, -8 and -9).

Of the many regulatory influences that modulate Fas-mediated apoptosis, regulation by FLIP stands out. Flip is found in several isoforms that are structurally similar to caspase-8 but lack caspase activity. Flip can be incorporated into Fas-DISCs, thereby preventing DISC-mediated processing and release of caspase-8.

15.6.2
Tumor Necrosis Factor-Receptor 1 and Apoptosis

The extracellular signaling protein tumor necrosis factor (TNF) is a major mediator of apoptosis and of inflammatory responses (reviews: Baud and Karin, 2001; Chen and Goeddel, 2002). By binding to cognate receptors, TNF-R1 or TNF-R2, several signal transduction pathways are activated. TNF-R1 activation mediates most of the biological activities of TNF. Binding of TNF to TNF-R1 triggers a series of cellular events, among which the activation of caspase-8 and the activation of two major transcription factors, NFκB and c-Jun stand out (Fig. 15.11).

The initial step in TNF-R1 activation involves the binding of the TNF trimer to TNF-R1, resulting in receptor clustering and release of an inhibitory protein (silencer of death domains, SODD) from TNF-R1's intracellular domain. Subsequently, the adaptor protein TRADD associates with the intracellular domain of the receptor and recruits additional adaptor proteins including FADD, which allows the binding and activation of caspase-8 within the TNF-R1 multiprotein complex.

The other adaptor proteins that are found in the TNF-R1 complex (review: Chen and Goeddel, 2002) mediate the recruitment and activation of protein kinases like the in-

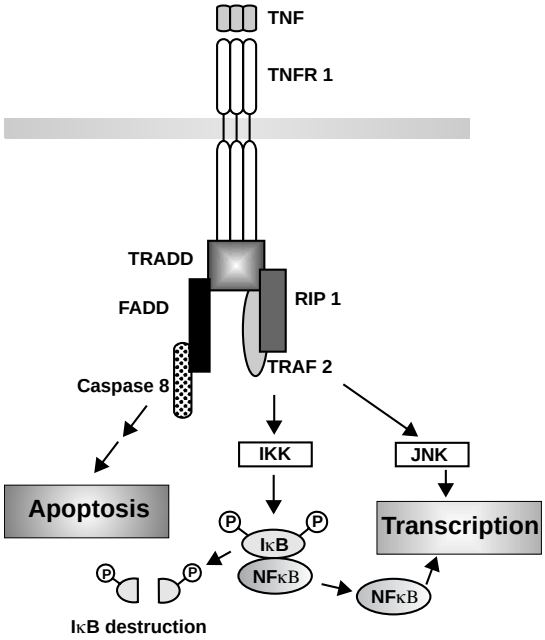


Fig. 15.11 Signaling by the tumor necrosis factor (TNF) receptor. Binding of TNF to its receptor induces association and activation for further signaling of several proteins which activate distinct signaling pathways. Assembly of the multiprotein complex on the cytoplasmic side is mediated mainly via death domains (DD) of the receptor and the adaptor protein TRADD. FADD induces apoptosis via activation of initiator caspase 8. TRAF2 and RIP mediate activation of transcription via two main ways. One way uses phosphorylation of the inhibitor IκB by IκB kinase (IKK) to induce its ubiquitin-mediated proteolytic destruction and the relieve of NFκB inhibition. Another way leads to activation of the JNK pathway (see Chapter 10) and stimulation of transcription of diverse target genes.

hibitor of NFκB kinases, IκB (see also Section 2.6.4), which link the TNF-signaling pathway to the NFκB function, providing for an anti-apoptotic and proliferation-promoting signal. Other signals from the activated TNF-R1 complex lead via MAPK pathways to the c-Jun terminal kinase, JNK, and to the activation of transcription factors including c-Jun (see Chapter 10).

15.7 Links of Apoptosis and Cellular Signaling Pathways

Like most functions in animal cells, the apoptotic program is regulated by signals from other cells, which can activate or suppress. In addition to these extracellular controls,

the apoptotic program is also controlled by intracellular signaling pathways. At different levels of the apoptotic program, there are links to cell-cell interactions, to growth-factor-controlled signaling pathways, to the cell cycle, and to the DNA damage checkpoint system. As discussed in Chapter 14, suppression of apoptosis is a crucial step in tumorigenesis, and numerous links exist between malfunction of apoptotic proteins and tumor formation.

Overall, our knowledge of links to intracellular and extracellular signaling pathways is very incomplete, and a detailed understanding is limited to a few examples.

Two examples are highlighted below.

15.7.1 PI3-Kinase/Akt Kinase and Apoptosis

The PI3-kinase/Akt kinase pathway (see Section 6.6) is an example of a signaling pathway that has a distinct anti-apoptotic function and promotes cell survival. It can mediate anti-apoptotic signals as well as growth-promoting signals (Fig. 14.12). The anti-apoptotic signal conduction starts at PI3-kinase to Akt kinase, which is activated by the messenger substance $PtdInsP_3$ formed by PI3-kinase. Two main ways have been iden-

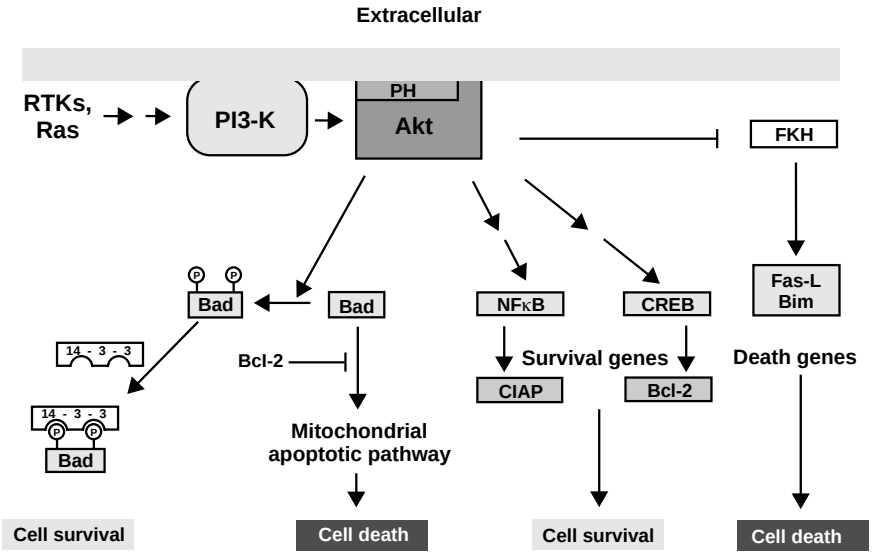


Fig. 15.12 Antiapoptotic signalling by the PI3-kinase/Akt kinase pathway. The PI3 kinase/Akt kinase pathway inhibits apoptosis and promotes cell survival via several ways. In one reaction, Akt kinase phosphorylates and inactivates Bad protein which is a proapoptotic protein. Phosphorylated Bad is bound by 14-3-3 proteins which makes it unavailable for triggering of apoptosis. Akt kinase also phosphorylates and activates the transcription factors NFκB and CREB which have the genes for the anti-apoptotic proteins IAP and Bcl-2 as targets. Members of the forkhead (FKH) family of transcription factors are inhibited upon phosphorylation by Akt kinase preventing transcription of the pro-apoptotic genes for Fas ligand and the Bim protein.

tified by which activated Akt kinase can influence the apoptotic program (review: Nicholson and Anderson, 2002).

In one way, Akt kinase promotes cell survival by directly phosphorylating transcription factors that control the expression of pro- and antiapoptotic genes. As an example, phosphorylation of proteins of the forkhead family of transcription factors by Akt kinase changes their subcellular localization. Forkhead proteins reside predominantly in the nucleus, where they activate transcription of proapoptotic target genes including *CD95 ligand* and *Bim*.

Activated Akt kinase phosphorylates forkhead proteins, leading to their export from the nucleus and sequestration in the cytoplasm by binding to 14-3-3 proteins. This negative regulation is contrasted by a positive regulation of the activity of the transcription factor NFκB, which is involved in the regulation of cell proliferation, apoptosis, and survival in response to a wide range of growth factors and cytokines. A large part of the survival-promoting function of NFκB is mediated through its ability to induce prosurvival genes such as IAP genes (see Section 15.3).

In a second way in which Akt kinase controls apoptosis, Akt kinase directly phosphorylates key regulators of apoptosis. The best-studied example of this type of control involves the Bad protein, which is a proapoptotic member of the Bcl-2 family. The Bad protein is phosphorylated by Akt kinase at Ser residues, and this modification promotes translocation of Bad to the cytosol, where it is found complexed with 14-3-3 proteins. By this mechanism, the pro-apoptotic effect of Bad can be inhibited. The effect on Bad is, however, not universal and is observed only in some cell types. Another pro-apoptotic substrate of Akt kinase is procaspase-9, which is inhibited upon phosphorylation by Akt. Dysregulation of the PI3-kinase/Akt kinase pathway, e. g., by inactivation of the PTEN tumor suppressor, has an anti-apoptotic effect and will favor tumor formation by preventing the death of cells that would be channeled to apoptosis under normal circumstances.

15.7.2 The Protein p53 and Apoptosis

The tumor suppressor protein p53 has both growth-inhibiting and pro-apoptotic properties that are essential to its tumor-suppressing activity. These functions of p53 can be separated and are mediated by distinct pathways. As outlined in Section 14.8.3, the growth-controlling activity is mediated mainly by the kinase inhibitor $p21^{CIP1}$, which is regulated by p53 at the level of expression. In addition, p53 can exert a pro-apoptotic function which is separate from the growth-inhibiting function. Apoptosis induced by p53 is especially important during conditions of DNA damage and stress. It can be categorized into transcription-dependent and transcription-independent reactions.

Apoptotic Genes Activated by p53

The list of genes activated by p53 includes many genes known to be important for apoptosis (review: Hickman et al., 2002). Examples (Fig. 15.13) include members of the Bcl-2 family of proteins, e. g., Bax, Bcl-2, Puma, Nora, death receptors and their

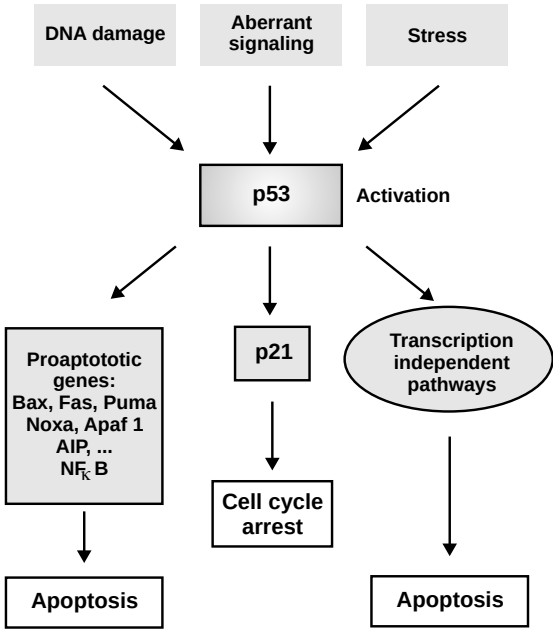


Fig. 15.13 Pathways of DNA damage-mediated and p53-mediated apoptosis. The tumor suppressor protein p53 is activated by DNA damage, malfunction of signaling pathways and by various stress influences. In a transcription-dependent pathway, p53 functions as a transcription activator of various proaoptotic genes which trigger the apoptotic program and lead to cell death. Furthermore, p53 activates transcription of the inhibitor $p21^{Kip1}$ leading to cell cycle arrest via inhibition of CDKs. We also know of less well characterized, transcription independent pathways (e. g. direct interaction with mitochondria) by which p53 can activate the apoptotic program.

ligands (Fas ligand, Fas/CD95, DR5), components of the mitochondrial path of apoptosis (Apaf1), and transcription factors (NFκB). Furthermore, p53 represses transcription of the anti-apoptotic protein Bcl-2. Although all of these proteins have been shown to be required for p53-mediated apoptosis in some cell systems, no single target gene has been identified as pivotal to the apoptotic pathway. It appears that the relative contribution of the p53-controlled proapoptotic genes to p53-mediated apoptosis is specific to the cell type. Depending on the cellular context, post-translational modifications of p53, e. g., phosphorylation or acetylation, may influence the expression pattern of apoptotic target genes. In the same sense, a cell-type specific interaction with distinct transcriptional cofactors will influence the choice of p53 target genes. A loss of p53 function is thought to change the levels of important proapoptotic proteins and to allow survival of damaged cells that would otherwise die by apoptosis.

Transcription-independent Induction of Apoptosis by p53

Transactivation of proapoptotic genes is not the only way that p53 protein can activate the apoptotic program. Non-transcriptional ways of p53-mediated apoptosis have also been described. As an example, a p53-regulated redistribution of the Fas receptor from the cytosol to the cell membrane has been demonstrated (Bennet et al., 1998). Furthermore, p53 can contribute to apoptosis by direct signaling to the mitochondria (Mihara et. al, 2003). These pathways, however, have been only incompletely characterized.

Reference

1 Baud, V. and Karin, M. (2001) Signal transduction by tumor necrosis factor and its relatives. *Trends Cell Biol.*, **11**, 372–377.

2 Bennett, M., Macdonald, K., Chan, S.W., Luzio, J.P., Simari, R. and Weissberg, P. (1998) Cell surface trafficking of Fas: a rapid mechanism of p53-mediated apoptosis. *Science*, **282**, 290–293.

3 Chen, G. and Goeddel, D.V. (2002) TNF-R1 signaling: a beautiful pathway. *Science*, **296**, 1634–1635.

4 Cory, S. and Adams, J.M. (2002) The Bcl2 family: regulators of the cellular life-or-death switch. *Nat.Rev.Cancer*, **2**, 647–656.

5 Enari, M., Sakahira, H., Yokoyama, H., Okawa, K., Iwamatsu, A. and Nagata, S. (1998) A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD. *Nature*, **391**, 43–50.

6 Evan, G. and Littlewood, T. (1998) A matter of life and cell death. *Science*, **281**, 1317–1322.

7 Hengartner, M. (1998) Apoptosis. Death by crowd control. *Science*, **281**, 1298–1299.

8 Hickman, E.S., Moroni, M.C. and Helin, K. (2002) The role of p53 and pRB in apoptosis and cancer. *Curr.Opin.Genet.Dev.*, **12**, 60–66.

9 Krammer, P.H. (2000) CD95's deadly mission in the immune system. *Nature*, **407**, 789–795.

10 Krammer, P.H. (2000) CD95's deadly mission in the immune system. *Nature*, **407**, 789–795.

11 Mihara, M., Erster, S., Zaika, A., Petrenko, O., Chittenden, T., Pancoska, P. and Moll, U.M. (2003) p53 Has a Direct Apoptogenic Role at the Mitochondria. *Mol.Cell*, **11**, 577–590.

12 Nicholson, K.M. and Anderson, N.G. (2002) The protein kinase B/Akt signalling pathway in human malignancy. *Cell Signal.*, **14**, 381–395.

13 Parone, P.A., James, D. and Martinou, J.C. (2002) Mitochondria: regulating the inevitable. *Biochimie*, **84**, 105–111.

14 Rathmell, J.C. and Thompson, C.B. (2002) Pathways of apoptosis in lymphocyte development, homeostasis, and disease. *Cell*, **109 Suppl**, S97–107.

15 Rathmell, J.C. and Thompson, C.B. (2002) Pathways of apoptosis in lymphocyte development, homeostasis, and disease. *Cell*, **109 Suppl**, S97–107.

16 Stennicke, H.R., Ryan, C.A. and Salvesen, G.S. (2002) Reprieval from execution: the molecular basis of caspase inhibition. *Trends Biochem.Sci.*, **27**, 94–101.

17 Vaux, D.L. and Korsmeyer, S.J. (1999) Cell death in development. *Cell*, **96**, 245–254.

18 Weber, C.H. and Vincenz, C. (2001) The death domain superfamily: a tale of two interfaces? *Trends Biochem.Sci.*, **26**, 475–481.

Index

A

abl gene 483
Abl
– c-Abl 342
– tyrosine kinase 333, 341
acetylcholine receptor 246
ACF 57
activation segment 275, 437
ADA 60
adaptor proteins 119, 121, 351
adenyl cyclase 206, 216, 220, 296
ADP
– cyclic ADP ribose 241, 243–244
– poly-ADP-ribose polymerase 44, 519
ADP-ribosylation 205
adrenaline 132, 137
β-adrenergic receptor 137
β-adrenergic receptor kinase (β-ARK) 194, 216, 272
AE-1 267
AF-1 156, 164
AF-2 156, 164
AF-6 379
agonists 131–132
AIF (apoptosis-inducing factor) 522
– AlF_4^- 210
AKAPs (A-kinase anchor proteins) 283, 290, 306–307
Akt
– kinase 249, 252–253, 288, 529
– pathway 83
aldosterone 129, 152
allostery 92
all-trans-retinoic acid 152
alprenolol 132
Alzheimer precursor protein (APP) 422
γ-aminobutyric acid receptors (GABA) 190
AMPA receptor 296
anaphase-promoting complex (APC) 108, 452, 507
antagonists 131–132
T antigen 493
AP1 24, 50
Apaf1 524
APC (anaphase-promoting complex) 108, 452, 507
apoptosis 511–531
– inhibitor of (IAP) 520
apoptosis-inducing factor (AIF) 210, 522
aporeceptor 171
APP (*Alzheimer* precursor protein) 422
ARAM 411
ARF protein 149, 491, 494, 504
– p19ARF 380
– subfamilies 355
Arg finger 366
β-ARK (β-adrenergic receptor kinase) 194, 216, 272
arrestin 195
ATF (activating transcription factor) 20, 28, 50
ATM (ataxia telangiectasia mutated) kinase 342, 466, 503
ATR (ataxia and Rad-related) kinase 466, 503
autocrine loops 481
autoinhibition 278
autoregulation 24
axin 508
5-azacytidine 66

B

Bad proteins 522
Bak proteins 521
Bax proteins 521, 530
Bcl-2 protein 482, 499, 517, 520, 530
Bcr protein 372
BH domain 521

Biochemistry of Signal Transduction and Regulation. 3rd Edition. Gerhard Krauss
Copyright © 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30591-2

bicoid 46
Bid proteins 517
Bim proteins 522, 530
BRCA genes 489
– BRCA1 106
Bub protein 475
– Bub1 gene 490
Burkitt's lymphoma 486

C

Ca^{2+} receptors 255
Ca^{2+}-ATPases 246
Ca^{2+} / calmodulin 223, 257–258, 279, 292
Ca^{2+}-myristoyl switch 149, 259
c-Abl 342
c-abl gene 483
E-cadherin 422
CAK (CDC2 activating kinase) 37, 437
calcineurin 258, 296, 302–304, 306
calcitonin
– genes 71
– receptors 191
calmodulin 95, 256–257
calreticulin 241
calsequestrin 241
CaM kinase
– CaM kinase II 219, 247, 292–293
– CaM kinase phosphatase 296
CaMKK 296
cAMP 137
– concentration 281
– phosphodiesterase 233, 282
cAMP-gated ion channels 233
cAMP-specific phosphodesterase 308
capping 37, 69
caretaker 487
caspase 515
– effector caspases 515
– ICAD (inhibitor of caspase-activated DNAase) 517
– initiator caspases 515
– pro-caspase 9 252
β-catenin 507
Cbl protein 108, 482
CBP protein 44, 60, 64
– CBP / p300 168
CD4 protein 411
CD8 protein 411
CD40 protein 396
CD45 protein 348
CD95 protein 525, 530
– CD95 / Fas 525
– CD95 ligand 530
CDC2 activating kinase (CAK) 37, 437
CDC6 protein 463
CDC7-Dbf4 kinase 463
CDC20 protein 453
CDC25 phosphatase 335, 343, 439, 456, 463–464
CDKs (cyclin-dependent protein kinases) 435–437
– CDK2 (cyclin A) 442, 460
– CDK2/4 (cyclin D) 454
– CDK5 436
– CDK7 (cyclin H) 37, 436
– CDK8 (cyclin C) 37, 436
– CDK9 (cyclin T) 37, 437
– CDK10 437
ceramide 261
cGMP 139, 235–236, 266
– phosphodiesterase 139, 208
cGMP-dependent protein kinases 235
cGMP-gated cation channels 236
cGMP-regulated protein kinase 272
C-helix 274
Chk1 466
Chk2 466, 503
cholera toxin 205
chromatin remodeling 57
CIP / KIP family 445
cis
– elements 17
– 9-*cis* retinoic acid 130, 152
c-jun 60, 528
C-kinase, activated, receptors for (RACK) 290, 307–308
CKIs (cyclin-dependent kinase inhibitors) 445, 493
c-Myc protein 481, 494
coactivators 42, 155
coexpressors 155, 168
colocalization 123–124
cortisol 129, 152
CpG islands 65
CRE protein 50–51
CREB protein 50–51, 60, 391
Crk protein 352
crosstalk 124–125
Csk 340
CSL 422
CTD (C-terminal domain) 36, 58
CTF / NF1 40
Cul1 111
cyclins 439–440, 442
– cyclin A (CDK2) 442, 460
– cyclin box 439
– cyclin C (CDK8) 37, 436
– cyclin D (CDK2/4) 454

– cyclin D1 155, 485
– cyclin E 454, 485
– cyclin H (CDK7) 37, 436
– cyclin T (CDK9) 37, 437
cyclin-dependent
– kinase inhibitors (CKIs) 445
– protein kinases (*see* CDKs) 37, 435–437, 442
cyclophylin 303
cyclosome 108
cyclosporin 303
cytochrome C 522
cytokine signaling, suppressor of (SOCS) 327

D

DAG (diacylglycerol) 225, 237, 259, 283
Dbl homology (DH) 371
death
– effector domain (DED) 520
– receptor 513, 520, 524
death-inducing signaling complex (DISC) 525
desensitization 192, 219
destruction box 453
DH (Dbl homology) 371
diacylglycerol (DAG) 225, 237, 259, 283
1,25-dihydroxycholecalciferol 129, 152
DISC (death-inducing signaling complex) 525
DlgA 336
DNA
– DNA damage checkpoints 466–467
– – G_1 DNA damage checkpoint 466–467
– – G_2 DNA damage checkpoint 467
– DNA methylation 472
– DNA methyltransferase 65
– DNA topoisomerase I 44
– E2F-DNA 342
– ICAD (inhibitor of caspase-activated DNAase) 517
– recognition elements 15
domain swapping 41
DREAM 23

E

E1 protein 102
– E1A protein 493
E2 protein 103
– E2F 62, 435, 457
– E2F-DNA 342
E3 protein 103, 105
E6 protein 110
– E6-AP 105, 110, 168
E7 protein 493
4E-BP 83
E-cadherin 422
EF structure 256
EF-TU 201
EGF (epidermal growth factor) 317
– receptor (EGFR) 109, 291, 327–328, 482
eIF-2 80–82
eIF-4 80, 83
– eIF-4E 392
Elk-1 390
endonuclease
– endonuclease G 522
– endonuclease HO 49
enhanceomes 63–65
enhancers 31
Epac 234
epidermal growth factor (*see* EGF) 109, 291, 317, 327–328, 482
epinephrine 130
ER (estrogen receptor) 166, 170
ErbB2 482
ErbB4 422
ERK 385
– ERK5 pathway 391
– pathway 388
erythropoietin 395
– receptor 398
estradiol 128, 152
estrogen receptor (ER) 166, 170

F

FAK (focal adhesion kinase) 415, 519
farnesylation 146
Fas protein 396
– CD95 / Fas 525
F-box proteins 105, 111, 451
feedback regulation 91
ferritin 76
– receptor (FGFR) 317
FHA (forkhead-associated) domains 336
fibroblast growth factor (*see* FGF) 317
FK506 303
FK506-binding protein 303
focal adhesion points 413
forkhead-associated (FHA) domains 336
forskolin 223
fos 50
Frizzled receptor 508
furin 422
Fyn 401
FYVE domain 335

G

G_1 DNA damage checkpoint 466
G_2 DNA damage checkpoint 467

G_{12} subfamily 204–205, 220
GABA (γ-aminobutyric acid receptors) 190
GADD45 498
GAL4 40, 42
GAL80 inhibitor 55
GAP (GTPase-activating proteins) 198, 356, 368–369
– GAP function 220
– Gap junctions 115
– p120 GAP 323, 368
GATA-1 24, 58
gatekeeper 487
Gbl 328
GCC2 kinase 81
GCKs (germinal center kinases) 392
GCN4 protein 8
GCN5 protein 44, 60
GDI (guanine nucleotide dissociation inhibitors) 199, 356–357
G-domain 212
GDP · AlF_4^- 210
GEF (guanine nucleotide exchange factors) 199, 206, 234, 356, 369–371
genetic imprinting 66
geranylation 146
germinal cneter kinases (GCKs) 392
G_i subfamily 204
glucagon receptor 190
glycine-rich loop 274
glycogen
– metabolism 299
– phosphorylation 97–99
– synthase kinase 272, 507
G_M subunit 299
gp130 398
GPI (glycosyl-phosphatidyl-inositol) anchor 147–148, 396
G-protein-coupled
– receptors 187–196
– – GRK (G-protein-coupled receptor protein kinases) 194
– signal transmission pathways 179–227
G_q subfamily 204
Grb2 protein 323, 331, 333, 352, 369
GRK (G-protein-coupled receptor protein kinases) 194, 333
growth hormone 395
G_s subfamily 204
GTP
– GTP analogues 200–201
– guanylyl cyclase from GTP 235
GTPases 121
– GTPase-activating proteins (*see* GAP) 16, 115, 198, 220, 323, 356, 368–369
– Ral GTPases 379
– Rho GTPases 220
guanine nucleotide
– dissociation inhibitors (GDI) 199, 356–357
– exchange factors (GEF) 199, 234, 356

H

H1 protein 56
H3 protein 56, 62, 68
H12 helix 163–164
HATs (histone acetyl transferase) 58
H-bonds 9
Hct1 453
HDAC (histone deacetylase) 58, 61–62, 68
heat shock protein (Hsp) 171
hect domain 105
helicase 35, 38
helix-loop-helix motif (HLH motif) 8, 50
7-helix transmembrane receptors 187
helix-turn-helix motif (HTH motif) 4
hemoglobin 266
heterodimers 17
heterotypic dimerization 50–52
heterozygosity, loss of (LOH) 487
hGH receptor 317
histamine 130
histone
– acetyl transferase (HATs) 58
– acetylase 34, 50, 58, 167, 495
– deacetylase (HDAC) 58, 61–62, 68, 168
– methylation 62
– octamer 55
HIV 73
HMG1 44, 58
– HMG1(Y) 44, 64
HMG2 44
hMSH2 489
HNPPC 489
HO endonuclease 49
hormone-responsive element (*see* HRE) 153, 157–159, 165, 172
hormones 128–129
– parat hormone receptors 190
– T_3 hormone 173
H-*Ras* 145
HRE (hormone-responsive element) 153, 157–159, 172
– negative HREs 165
– positive HREs 165
HRI 81
Hsp (heat shock protein) 171

I

IκB 23, 111
IAP (inhibitor of apoptosis) 520
ICAD (inhibitor of caspase-activated DNAase) 517
IFN (*see* interferon) 57, 63, 395
IL (*see* interleukins) 395, 398–399
immunoproteasome 108
InaD protein 290, 336
initiation element 30
initiator caspases 515
INK4 family 445, 447
inositol-1,4,5-triphosphate 225
[Ins(1,4,5)P_3] 237
InsP_3 receptor 241
insulin 83, 252
– receptor 313, 320
– – insulin receptor substrate (IRS) 320, 352–353
integrin 308, 413
interferon (IFN) 395
– IFNβ 57, 63
– IFN regulary factor 63
interleukins (ILs) 395
– IL-2 receptor 298–399
– IL-6 receptor 398
ion-channel
– lignand-gated 181
– voltage-gated 181
IRA1 368
IRA2 368
IREs (iron-response elements) 77
IRPs (iron regulatory proteins) 77
– IRP1 79
IRS (insulin receptor substrate) 320, 352–353
isocytrate dehydrogenase 100
isoprenylation 146
isoproterenol 132
ITAM 411

J

Jak kinases 405
Jak-Stat pathway 405
Janus kinase 401, 425
JIP1s 387
JNK 528
– JNK / SAPK pathway 391
jun 50
– *c-jun* 60, 528

K

kinetic bilayer trapping 145
K-*Ras* 144

L

Lac repressor 15
LAT (linker for activation of T-lymphocytes) 411
Lck 401
leaky scanning 78
leucin zipper 8, 50
leukemia inhibitory factor (LIF) 398
lexA repressor 25
LIF (leukemia inhibitory factor) 398
ligand-gated ion-channel 181
lipid anchors 143–144
LOH (loss of heterozygosity) 487
loop
– autocrine loops 481
– helix-loop-helix motif (HLH motif) 8, 50
– P-loop 201, 274, 361
– T-loop 437
LPA (lysophosphatidic acid) 261
L-triiodothyronine 152
Lyn 288
lysophosphatidic acid (LPA) 261

M

MAP kinase (MAPK) 296, 375
– cascade 195
– MAP4 (MAP4Ks) 386, 448
– module 384
– pathways 383–393
MAPKAP-1 kinase 391
MAPKK kinase 385
MARCKS 149, 290
Max protein 8, 485
MBD1 66
MBD2 66
MBD3 66
MBD4 66
MCLK (myosin light chain kinase) 258, 272, 292
MCM proteins 448, 461
MDM2 protein 106, 459, 500–501
MdmX 504
MeCP2 66
mediators 36, 43, 155
MEK
– kinase 378, 385–386
– proteins 385
membrane
– anchoring 142–149
– receptors with associated tyrosine kinase activity 395–416
5-methyl cytidine 65
N-methyl-D-asparte receptors 307
methylation 24

MHC complex 409
mixed lineage kinases (MLKs) 392
MNK 392
MOS kinase 388
MP1 387
mRNA (*see also* RNA)
– processing 37
– stability 75
mSin3 168
mSos protein 370
mTOR-pathway 83
mutator phenotype 474
Myc
– c-Myc protein 481, 494
– Myc transcription 485
myosin light chain kinase (MCLK) 258, 272, 292
myristoyl, Ca^{2+}-myristoyl switch 149
myristoyl-electrostatic switches 149
myristoyl-ligand switches 149
myristoylation 144

N

Na^{+}-Ca^{2+} exchange proteins 246
NAADP (nicotinic acid adenine dinucleotide phosphate) 241, 244
NC2 52, 55
Nck 333
NcoR 168
Nef 341
N-end rule enzymes 108
NER (nucleotide excision repair) 38
neurofibromin 368
NFκB 8, 23, 63, 111, 113, 170, 247, 391, 528
NF-AT 247, 303
NGFI-B 159
nicotinic acid adenine dinucleotide phosphate (NAADP) 241, 244
nitrosative stress 264
S-nitrosylation 263
NMDA receptor 264, 291, 296, 377
NO 133, 235
– guanylyl cyclase, NO-sensitive 266
– NO signaling molecule 261–267
– NO synthase 263, 296, 377
Nora proteins 530
noradrenaline 137
Nore1 379
norepinephrine 130
NOTCH protein 422
Noxa proteins 522
NTF-1 43
nuclear
– localization sequences 48
– signaling by nuclear receptors 151–176
nucleosome 55
nucleotide
– exchange factor (GEF) 206, 369–371
– excision repair (NER) 38
NuRD complex 57, 168

O

OmpF 184
ORC (origin recognition complex) 461
orphan receptors 156

P

P13-kinase 216, 248–253, 323, 325, 332, 379, 403, 484, 529
P13-like kinase 466
p19ARF 380
p21-activated protein kinases (PAKs) 392, 519
– $p21^{CIP1}$ 447
– $p27^{KIP1}$ 446
p38 pathway 391
p53 58, 105, 110, 113, 460, 491, 494–505
p63 503
$p70^{56}$ kinase 85
p73 504
p107 447, 456
p120 GAP 323, 368
p130 447, 456
p300 44, 60, 502
– CBP / p300 168
palindromic structure 158
palmitoylation 145
parat hormone receptors 190
paxillin 415
PCAF 43, 58, 60, 64, 167, 495, 502
PDGF receptor 125, 250, 318, 328, 341
PDK1 (phosphoinositide-dependent protein kinase 1) 252, 288
PDZ domains 227, 296, 336, 353
PERK 81
pertussis toxin 206
PEST sequences 109, 452
PH-(*Pleckstrin* homology)-domain 195, 225, 251, 334
Philadelphia translocation 483
6-phosphofructo-2-kinase 252
phorbol esters 284
phosducin 218
phosphodicin 216
phosphoinositide-dependent protein kinase 1 (PDK1) 252, 288
phospholipase
– phospholipase A2 254, 391
– phospholipase C 216, 225–227, 237

– – phospholipase Cβ 208, 247, 283
– – phospholipase Cγ 283, 323, 326, 331
phosphoprotein-ubiquitin ligase 109
phosphorylation 24
– glycogen 97–99
– protein 96
phosphotyrosine-binding (PTB) domain 332
Pin1 335
PKR 81
PL-Cβ 237
PL-Cγ 237
Pleckstrin homology-(PH)-domain 195, 225, 251, 334
P-loop 201, 274, 361
polyadenylation 37, 69–71
– alternative 70
poly-ADP-ribose polymerase 44, 519
postsynaptic densities (*see* PSD) 296, 336, 353
PP1 (protein phosphatase 1) 299–301
PP2 (protein phosphatase 2) 301–304, 335
– PP2A 301–302, 335
– PP2B 302–304
PPAR 157, 163
practolol 132
pRB protein (retinoblastoma protein) 62, 342, 447, 456–460, 490–493
pre-initiation complex 32
presenilin 422
procaspase 9 252
progesterone 128, 152
propranolol 132
pro-rich domains 249
prostaglandine
– prostaglandine E_2 130
– prostaglandine J_2 152
protease-activated receptors 187
proteasome 107–108
– immunoproteasome 108
protein
– 14-3-3 protein 335, 375, 467, 530
– 14-3-3 σ protein 498
– cyclin-dependent protein kinases (*see* CDKs) 37, 435–437
– G-protein-coupled receptors (*see there*) 187–196
– kinases 120
– – protein kinase A (PKA) 193, 233–234, 272, 280
– – protein kinase B (PKB) 252
– – protein kinase C (PKC) 95, 193, 223, 255, 260, 272, 283, 336, 375
– modules 122–123
– phosphatases 296
– – protein tyrosine phosphatases 342–351
– – protein phosphatase I (PP I) 299–301
– – protein phosphatase 2A (PP 2A) 301–302
– – protein phosphatase 2B (PP 2B) 302–304
– phosphates 120
– phosphorylation 96
PSD (postsynaptic densities) 296
– PSD-95 336, 353
PTB (phosphotyrosine-binding) domain 332
PtdIns(3,4,5)P_3 249
PtdIns(4,5)P_2 225, 249
PTEN phosphatase 250, 253, 530
PTPs 343
– SH-PTP1 343, 349
– SH-PTP2 323, 343, 349
Puma proteins 522, 530
pyrin domain 520

R

Rab protein family 355, 357–358
RACK receptors for activated C-kinase 290, 307–308
Raf kinase 122, 291, 373–375, 480
Ral
– Ral GDS 379
– Ral GTPases 379
Ran protein / protein family 355, 358
– RanGAP1 113
Rap1 234
– Rap1A 374
Rap46 173
Rar 157, 164, 176
Ras protein / *Ras* protein family 144–145, 199, 250, 355–380, 480, 483
– H-*Ras* 145
– K-*Ras* 144
– *Ras* pathway 326
*Ras*GRF1 377
*Ras*GRP 377
receptors for activated C-kinase (RACK) 290, 307–308
recoverin 149, 258
redox-regulation 24
regulated intramembrane proteolysis (RIP) 422
repressors 28
– Lac 15
– lexA 25
– Trp (translation repressor) 21, 79, 336
– URS (upstream repressing sequences) 31
response regulator 425
retinal ligand 191
retinoblastoma protein pRb 62, 342, 447, 456–460, 490–493
retinoids 130, 173

– 9-*cis* retinoic acid 130, 152
Rev protein 73
– Rev responsive element 75
Rgl 379
RGS (regulators of G protein) 208, 212, 216, 218, 222, 247
– RGS9 199
Rho protein family 357
– Rho GTPases 220
– Rho / Rac 355
rhodopsin 183, 188, 192
Rin1 379
RING finger , 328
RIP (regulated intramembrane proteolysis) 422
RNA polymerase
– of *E. coli* 26
– holoenzyme 26, 28
– RNA polymerase II 30, 35
– RNA polymerase III 30
RSK 391
RXR 157, 173–176
– heterodimer 157, 173, 175
RXR-T_3R heterodimer 162
ryanodin receptors 241
RZR 159

S

SAGA 43, 60
SAPK 387
Sar / Arf protein family 358
SCF complex 109, 451–452
second messengers 121
γ-secretase 422
SH2 domain 249, 320, 329–332
SH3 domain 249, 332–334
Shc 323, 331
– adaptor protein 403
Shine-Dalgarno sequence 78
SH-PTP1 343, 349
SH-PTP2 323, 343, 349
σ-factor 26
signal
– amplification 139
– pathways 115–149
– recognition particle (SRP) 202
silencers 32
Skp1 111
SLN1 425
Smac / Diablo proteins 522
SMAD-proteins 49, 418–421
SMRT 168
snRNPs (small nuclear ribonucleoproteins) 71
SOCS (suppressor of cytokine signaling) proteins 327, 404
SODD 527
Sp1 43
spliceosome 71
splicing, alternative 70–71
SR proteins 37
Srb / mediator complex 37
SRB proteins 36, 43
Src
– gene 478
– kinase 288, 308, 323, 332, 338–341, 348
– – tyrosine kinase 329
SREBP (sterol regulatory eleent-binding protein) 422
SRP (signal recognition particle) 202
SSK1 425
Stat
– factors 60
– Jak-Stat pathway 405
– proteins 49, 331, 403, 406
Ste5 387
sterol regulatory element-binding protein (SREBP) 422
SUMO 113
suppressor of cytokine signaling (SOCS) proteins 327, 404
SUV39 62
SW1 / SNF 57
SW15 49
SWi / SNF 168
Swi5 60
switch
– Ca^{2+}-myristoyl switch 149
– myristoyl-electrostatic switches 149
– myristoyl-ligand switches 149
– switch I 361
– switch II 361
Syk 415

T

T antigen 493
T cell antigen receptor 410
T_3 hormone 129, 173
T_3 receptor (T_3R) 157, 165
TACE 422
TAFs 34, 43, 55–56
– $TAF_{II}250$ 34, 43, 56, 60
TAOs (thousand and one kinases) 392
Tat protein 73
TATA box 30, 54
– TATA box-binding protein 11–13
TBP 34, 52, 54–55, 65
testosterone 129, 152

tetradecanoyl phorbol acetate (TPA) 284
TFIIA 35
TFIIB 35, 54
TFIID 34
TFIIE 35
TFIIF 35
TFIIH 35, 38
TFIIIA 161
TFIIIB 448
TFR (transferrin receptor) 76
TGFβ
– family 417–418, 456, 460
– receptor 272, 417–418
thrombin 187
thrombospodin-1 gene 499
T-loop 437
TNF (tumor necrosis factor) 111, 396
– receptor (TNFR) 392
– TNFα 513, 524, 527
– TNF-R1 524, 527
mTOR-pathway 83
TPA (tetradecanoyl phorbol acetate) 284
TRADD 527
transcortin 171
transcription / transcriptional
– ATF (activating transcription factor) 20, 28, 50
– general transcription factors 33
– regulation of 1–2
transduceomes 305
transducin 139, 199
transferrin receptor (TFR) 76
transition state analog 212
translation
– factor 80
– repressor (Trp) 21, 79, 336
transmembrane
– domain 183–185
– elements 183–184
TRAP 167
L-triiodothyronine 152
tropomyosin 72
troponin 72
– troponin C 258
Trp repressor 21, 79, 336
– Trp channels 272
tubulin 76
tumor necrosis factor (*see* TNF) 111, 392, 396, 513, 524, 527
Tyk 401
tyrosine
– Abl tyrosine kinase 333, 341
– PTB (phosphotyrosine-binding) domain 332
– Src kinase 329
– tyrosine kinase activity 395–416

U

UAS (upstream activating sequences) 31
upiquitin 102–109
– pathway 440, 450
– phosphoprotein-ubiquitin ligase 109
ubiquitin-proteasome 108–109
upiquitin-protein-ligase 103
ubiquitination 170
upstream
– activating sequences (UAS) 31
– repressing sequences (URS) 31
URS (upstream repressing sequences) 31

V

van der Waals contacts 11
Vav oncoprotein 372
VDR 157
visual process 139
vitamin D3 173
voltage-gated ion-channel 181

W

WD
– WD motifs 308
– WD repeat 215
Wee-1 kinase 437
Wnt 507
WSXWS motif 396
WW domains 335

X

XPB 38
XPD 38

Y

Yotiao 307

Z

Zap 70 109
– kinase 411
Zif268 5, 161
zinc binding motifs 5
Zn-motifs 160
ZO-1 336